Searches for Direct Pair Production of Scalar Tau Leptons in Events with Two Hadronically Decaying Tau Leptons with the ATLAS Detector in LHC Run 2

Clara Elisabeth Leitgeb

München 2020

Searches for Direct Pair Production of Scalar Tau Leptons in Events with Two Hadronically Decaying Tau Leptons with the ATLAS Detector in LHC Run 2

Dissertation an der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Clara Elisabeth Leitgeb
geboren in Salzburg

München, den 5.11.2020

Erstgutachter: PD. Dr. Alexander Mann
Zweitgutachter: Prof. Dr. Thomas Kuhr
Mündliche Prüfung: 18.12.2020

Zusammenfassung

Bis heute konnte aus experimentellen Befunden kein Widerspruch zum Standardmodell der Teilchenphysik abgeleitet werden. Es gibt jedoch einige Hinweise für die Existenz einer fundamentaleren zugrundeliegenden Theorie bei höheren Energieskalen. Am Large Hadron Collider kann das mögliche Auftreten von neuer Physik in Proton-Proton Kollisionen bei zuvor unerreichten Schwerpunktsenergien von derzeit 13 TeV erforscht werden. Einige der vielversprechendsten Erweiterungen des Standardmodells sind supersymmetrische Modelle, die ein gänzlich neues Spektrum von Teilchen vorhersagen. Diese werden durch eine Symmetrie von Fermionen und Bosonen mit den Teilchen des Standardmodells in Verbindung gebracht.

In dieser Dissertation werden zwei Ansätze für die Suche nach dem supersymmetrischen Partnerteilchen des τ-Leptons im Standardmodell (Stau oder $\tilde{\tau}$) mit $139\,\text{fb}^{-1}$ an Daten vorgestellt, die durch den ATLAS-Detektor am LHC aufgezeichnet wurden. Die selektierten Kollisionsereignisse beinhalten zwei hadronisch zerfallende τ-Leptonen, sowie fehlenden transversalen Impuls. Der erste Ansatz besteht in einer Analyse, die auf Schnitten auf kinematische Variablen basiert. Im Zuge dieser Analyse konnten – in Abwesenheit eines signifikanten Überschusses in den Daten – Staumassen zwischen 120 GeV und 390 GeV mit einem beinahe masselosen leichtesten supersymmetrischen Teilchen (LSP) bei 95% Konfidenzniveau ausgeschlossen werden. Des Weiteren war es möglich, die Paarproduktion des supersymmetrischen Partners des linkshändigen τ-Leptons ($\tilde{\tau}_L$) gesondert im Staumassen-Bereich von 155 GeV bis 310 GeV auf einem Konfidenzniveau von 95% auszuschließen. Für den supersymmetrischen Partner des rechtshändigen τ-Leptons ($\tilde{\tau}_R$) konnte keine Sensitivität erreicht werden, die einen Ausschluss von Staumassen unabhängig von der Paarproduktion von $\tilde{\tau}_L$-Teilchen ermöglicht hätte.

Die zweite Suche hat zum Ziel, die Sensitivität der ersten Analyse zu verbessern, indem Boosted Decision Trees und ein Multi-Bin-Fit verwendet werden. Die erwartete Ausschluss-Sensitivität für Staumassen reicht von 80 GeV bis 430 GeV für die kombinierte $\tilde{\tau}_L$- und $\tilde{\tau}_R$-Paarproduktion. Für die gesonderte Produktion von $\tilde{\tau}_L$-Paaren wird ein Ausschluss von Staumassen zwischen 100 GeV und 350 GeV erwartet. Weiterhin konnte auch für eine Spanne von Staumassen zwischen 120 GeV und 230 GeV für gesonderte $\tilde{\tau}_R$-Produktion eine erwartete Sensitivität erreicht werden, die einen Ausschluss dieser Massen ermöglichen kann.

Abstract

The Standard Model of particle physics has so far not been contradicted by evidence from experiments. There are, however, several indications that a more fundamental underlying theory is necessary at higher energy scales. At the Large Hadron Collider the possible presence of new physics can be probed in proton-proton collisions at an unprecedented center-of-mass energy of 13 TeV. Some of the most promising extensions to the Standard Model are suggested by supersymmetric models, which predict an entirely new particle spectrum connected to the Standard Model particle content via a fermion-boson symmetry.

In this thesis, two approaches for the search for the supersymmetric partner particle to the Standard Model τ-lepton (stau or $\tilde{\tau}$) with $139\,\text{fb}^{-1}$ of data taken by the ATLAS detector at the LHC are presented. The considered collision events should contain two hadronically decaying τ-leptons and missing transverse momentum. The first approach is an analysis based on cuts on kinematic variables, which – in the absence of a supersymmetric signal in data – could exclude stau masses between 120 GeV and 390 GeV at 95% CL for a nearly massless lightest supersymmetric particle (LSP). It was also possible to exclude pair production of the supersymmetric partner of the left-handed τ-lepton ($\tilde{\tau}_L$) for stau masses between 155 GeV and 310 GeV at 95% CL. For the supersymmetric partner of the right-handed τ-lepton it was not possible to reach exclusion sensitivity separately from $\tilde{\tau}_L$ production.

The second search aims to improve the sensitivity of the first iteration of the analysis by using boosted decision trees and a shape fit. The expected exclusion sensitivity ranges from stau masses of 80 GeV to 430 GeV for the combined scenarios of $\tilde{\tau}_L$ and $\tilde{\tau}_R$ production. For the separate $\tilde{\tau}_L$ pair production, the expected exclusion range covers stau masses from 100 GeV to 350 GeV. Furthermore, for a range of stau masses between 120 GeV and 230 GeV exclusion sensitivity could also be achieved for separate $\tilde{\tau}_R$ pair production.

It is an old maxime of mine that when you have excluded the impossible, whatever remains, however improbable, must be the truth.

– Sherlock Holmes

Contents

1. Introduction

Centuries of experimental and theoretical research have shaped our understanding of the world around us. Quantitative models have been used to describe the known phenomena and predict new ones. Over time, models have been developed, refined, discarded and replaced by new ones, according to the experimental evidence. One of the most successful physical models is the current Standard Model of particle physics. Experimental tests have shown that its predictions are accurate to a very high precision. Nevertheless, there are strong indications for the presence of phenomena that the Standard Model in its current form cannot explain. From the experimental side, e.g. gravitational effects additionally to those that can be associated to "Standard Model-like" matter have been observed in the universe. There are also theoretical arguments as to why the current Standard Model cannot be the end of the story, for instance the hierarchy problem.

Theoretical particle physics provides numerous solutions to some of the shortcomings of the Standard Model. One of the most promising class of models are supersymmetric models, predicting the existence of a partner particle for each Standard Model particle, which differs in its spin by $\frac{1}{2}$. There are several ways to verify or falsify the existence of Supersymmetry in nature. One of them is to search for the additional particles it predicts in highly energetic collisions of protons, as they are done at the LHC at CERN.

This thesis presents two approaches to the search for the supersymmetric partner particle of the Standard Model τ-lepton ("stau") using the data collected by the ATLAS detector in 2015-2018 at the LHC. In both analyses the direct pair production of staus is considered to occur with a subsequent decay into one Standard Model τ-lepton and a lightest neutralino each. The lightest neutralino is assumed to be the lightest supersymmetric particle and since R-parity conservation is assumed, it is stable. Furthermore, the event selection for these searches focuses on the hadronic decay modes of the τ-leptons and an imbalance in the reconstructed transverse momenta, which is caused by the neutralinos escaping direct measurement in the detector.

Chapter 2 gives an overview of the theoretical foundations of the Standard Model, Supersymmetry and particularly the considered model for the stau search. The experimental setup of the LHC and the ATLAS detector are described in chapter 3, before going into some detail on the data that was used in the analyses, as well as the simulations and the reconstruction of physical objects based on detector signatures in chapter 4. After outlining the basic analysis strategies and some other general aspects of the search for stau pair production in chapter 5, the two searches for stau pair production are described in chapter 6 and chapter 7, respectively.

Throughout this thesis natural units ($\hbar = c = 1$) will be used.

2. Theoretical Foundations

2.1. Standard Model of Particle Physics

The Standard Model of particle physics is a framework of quantum field theories, describing all known elementary particles and their interactions. The series of discoveries of elementary particles stated in 1897, when the electron was experimentally confirmed by J.J. Thomson and his colleagues in experiments with cathode rays. The last elementary particle to be incorporated into the Standard Model was the Higgs boson which was found in 2012 at the LHC in proton-proton collisions. Since then no additional elementary particles have been discovered although it is known that the Standard Model in its current form cannot explain all the observed phenomena, e.g. in Astrophysics.

2.1.1. Free Spin 0, Spin $\frac{1}{2}$ and Spin 1 Particles

In classical mechanics each particle with momentum $\vec{p}$, energy E and mass m inside a potential V has to obey the classical energy and momentum relation:

$$\frac{\vec{p}}{2m} + V = E. \tag{2.1}$$

In a step that is known as "first quantization" the momentum and the energy of the particle can be replaced by the operators of the space and time derivatives acting on a wave function Ψ:

$$\vec{p} \to -i\nabla \tag{2.2}$$

$$E \to i\frac{\partial}{\partial t}. \tag{2.3}$$

This replacement results directly in the Schrödinger equation:

$$-\frac{1}{2m}\nabla^2\Psi + V\Psi = i\frac{\partial\Psi}{\partial t}. \tag{2.4}$$

However, the Schrödinger equation describes particles in non-relativistic quantum mechanics. If the description should also hold for relativistic particles, the relativistic energy-momentum relation must be obeyed ($c = \hbar = 1$, this time considering a free particle, so $V = 0$):

$$E^2 - \vec{p}^2 = m^2, \tag{2.5}$$

or equivalently using the Einstein summation convention:

$$p^\mu p_\mu - m^2 = 0. \tag{2.6}$$

Again replacing the four-momentum p^μ by

$$p^\mu \to i\partial_\mu, \tag{2.7}$$

the equation for a freely propagating spin 0 particle is obtained, the Klein-Gordon equation:

$$\left[\left(\frac{\partial}{\partial t}\right)^2 - \vec{\nabla}^2 + m^2\right]\psi(\vec{x}) = 0. \tag{2.8}$$

Historically, the Klein-Gordon equation lead to an interpretation issue as it allows for negative values of $|\psi(x)|^2$, which was understood as the probability density and thus had to be positive definite. To circumvent this issue, Paul Dirac thought of an equation which is first order in the time derivative instead, and which also obeys the relativistic energy-momentum relation:

$$E\psi = (\alpha m + \vec{\beta}\vec{p})\psi, \tag{2.9}$$

where α and $\vec{\beta}$ are constants which are constrained by the relativistic energy-momentum relation. A set of matrices that satisfy these constraints are the γ matrices:

$$\alpha = \gamma_0 = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \tag{2.10}$$

$$\alpha\beta_i = \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}. \tag{2.11}$$

Here, $\mathbb{I}$ denotes the 2×2 unit matrix and σ_i $(i = 1, 2, 3)$ the three Pauli matrices. Once again substituting $p_\mu = (E, \vec{p})$ with $i\partial_\mu$ the Dirac equation reads:

$$(i\gamma^\mu\partial_\mu - m)\psi = 0, \tag{2.12}$$

where ψ is now called a Dirac spinor and has four components. There are four solutions to this equation which are identified as spin $\frac{1}{2}$ fermions ($p^\mu = (E, \vec{p})$) and anti-fermions ($p^\mu = (-E, -\vec{p})$) with different spin orientations.

Finally, in relativistic quantum mechanics, a spin 1 field A^ν is characterized by the Proca equation:

$$\partial_\mu F^{\mu\nu} + m^2 A^\nu = 0, \tag{2.13}$$

where $F^{\mu\nu} \coloneqq \partial^\mu A^\nu - \partial^\nu A^\mu$ is the field strength tensor. The Standard Model in its current form only includes fundamental particles with spin $\frac{1}{2}$ (fermions), spin 1 (vector bosons) and spin 0 (the Higgs boson).

According to the kinds of interactions the fermions are subject to they can be further classified as quarks (interacting via the electroweak and strong force) and leptons (only interacting electroweakly). Quarks as well as leptons appear in three generations. The second and the third generation can be understood as a "copy" of the first one, which only differ by their larger particle masses (and consequently their lifetime) but not in any other property. An overview of the fundamental fermionic particles in the Standard Model is given in Tab.2.1.

Generation	Quarks	Mass [GeV]	Leptons	Mass [GeV]
1	up quark	$3 \cdot 10^{-3}$	electron neutrino	$< 10^{-9}$
	down quark	$5 \cdot 10^{-3}$	electron	$5 \cdot 10^{-4}$
2	charm quark	1.3	muon neutrino	$< 10^{-9}$
	strange quark	0.1	muon	0.106
3	top quark	173	τ-neutrino	$< 10^{-9}$
	bottom quark	4.5	τ-lepton	1.78

Table 2.1.: List of fundamental fermionic particles in the Standard Model with their respective (experimentally found) masses. [1]

The vector bosons arise in the course of establishing the interactions between particles by requiring local gauge invariance, which will be discussed in the following [2, 3].

2.1.2. Introducing Interactions: Gauge Theories

Considering the already introduced equation of motion for a free relativistic spin $\frac{1}{2}$ particle, it is evident that applying an arbitrary rotation by a constant angle θ to the Dirac spinor as

$$\psi(x) \to \psi'(x) = e^{iq\theta}\psi(x) \tag{2.14}$$

does not change the form of the equation. This is referred to as global gauge invariance. However, if the rotation angle is assumed to depend on the coordinates in space time, i.e. $\theta \to \theta(x)$, the Dirac equation's form is not preserved anymore, as terms dependent on $\partial_\mu\theta(x)$ arise which do not cancel out. To restore the invariance of the equation under what is called local gauge transformations, a new field has to be introduced which compensates for the non-vanishing terms. The gauge field also obeys a transformation rule such that the equations of motion for the particle, the new field and the interaction of both are invariant under local gauge transformations. The principle of local gauge invariance is used to introduce the electromagnetic, the electroweak and the strong gauge fields and therefore builds the foundations of the Standard Model of particle physics.

Quantum Electrodynamics

Quantum Electrodynamics was the first quantum field theory to be developed. Many people were involved in this process, among them R. Feynman, S. Tomonaga and J. Schwinger who were awarded with the Nobel Prize in 1965 for their efforts [4–8].

Continuing the discussion above, the Dirac equation preserves its form if the derivative ∂_μ

is replaced by a so-called covariant derivative:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iqA_\mu(x), \tag{2.15}$$

where a field A_μ has been introduced, which can be identified as the electromagnetic field. Under a local $U(1)$ rotation by $\theta(x)$, A_μ transforms as

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) - \partial_\mu\theta(x). \tag{2.16}$$

The total local gauge invariant Lagrangian now consists of three parts: The free propagation of the fermion with electric charge q and mass m, the interaction term of this fermion with the field A_μ and the free A_μ term. After writing out the covariant derivative it reads

$$\mathcal{L}_{\text{QED}} = \underbrace{[i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi]}_{\text{free fermion}} - \underbrace{(q\bar{\psi}\gamma^\mu\psi)A_\mu}_{\text{interaction term}} - \mathcal{L}_{\text{free field}}. \tag{2.17}$$

A_μ is a spin 1 field and thus the Proca Lagrangian for its free part is

$$\mathcal{L}_{\text{free field}} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}M^2A_\mu A^\mu, \tag{2.18}$$

with the field strength tensor $F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$. In order for this part of the total Lagrangian to be invariant under local U(1) transformations too, the mass term has to vanish, thus resulting in a massless photon [2, 9]:

$$\mathcal{L}_{\text{QED}} = \underbrace{[i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi]}_{\text{free fermion}} - \underbrace{(q\bar{\psi}\gamma^\mu\psi)A_\mu}_{\text{interaction term}} - \underbrace{\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}}_{\text{free photon}}. \tag{2.19}$$

This Lagrangian can be used to derive the form of the external lines, vertex factors and propagators in any QED Feynman diagram. Via the Feynman rules matrix elements and subsequently cross sections and decay widths can be calculated for QED reactions. However, as soon as higher-order diagrams, e.g. vacuum polarization diagrams containing fermion loops, are considered, divergences may appear in the matrix element calculation. For a long time this was a strong reason to doubt the validity of quantum field theories in general. However, today, these divergences are dealt with by means of regularization and renormalization. It has to be stated that only a very small fraction of theories is renormalizable at all – QED is one of them (as well as the spontaneously broken electroweak theory and QCD, which will be discussed in the following sections). First, the regularization step is performed, which replaces all infinities in the integration boundaries with some cut-off values, such that the integration itself can be performed in a meaningful way. Later on, these cut-offs can be taken to infinity again. Another possibility is the so-called dimensional regularization, where the integral is performed in $4 - 2\epsilon$ dimensions and ϵ is taken to 0 afterwards. Next, in the renormalization step the bare coupling constants and masses, which are inaccessible to measurements, are assumed to be divergent and cancel against the divergences arising from the Feynman diagrams, thus delivering finite results again. This redefinition of coupling constants, i.e. the electron charge, leads to a dependence on momentum transfer, which is known as "running couplings" and described by the Renormalization Group Equations (RGE) [10–12].

Weak Interactions

The first attempt for describing weak interactions was already made in 1934, when E. Fermi introduced a theory of point-like interactions involving four fermions with a universal coupling constant G_F [13]. Two decades later in 1957 it was suggested that weak interactions violate parity, which was also proven by experiment [14]. Weak interaction thus should work as a vector-like interaction which only affects the components of left-handed chirality of fermions and the right-handed chirality component of anti-fermions. This fact was implemented into the Fermi theory through the V-A structure of weak interactions. Furthermore, in systems of neutral K-mesons it was discovered that weak interactions not only violate parity but also the combined parity and charge symmetries – the so-called CP-violation. This was added into the theory of weak interactions as a complex phase in the Cabbibo-Kobayashi-Maskawa matrix (CKM matrix), thereby also introducing three generations of quarks [15, 16]. The 3×3 unitary CKM matrix transforms the strong force eigenstates of down-type quarks (d) into the slightly different weak eigenstates (d'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.20}$$

With the V-A structure of weak interactions and the complex phase of the CKM matrix providing maximal parity violation and CP-violation, respectively, the Fermi theory continued to be a good description of weak processes for phenomena at low energies. However, increasing the energy, the validity of the Fermi theory is limited by a unitarity bound. If a point-like scattering process involving four fermions is considered, the cross section is given as

$$\sigma = \frac{G_F^2}{\pi} \cdot s = \frac{G_F^2}{\pi} \cdot 4p^{*2}, \tag{2.21}$$

with $\sqrt{s} = 2p^*$ being the center-of-mass energy and p^* the momentum of one of the involved particles in the center-of-mass frame. In quantum mechanics, the cross section of inelastic scattering processes is derived by a partial wave expansion:

$$\sigma = \frac{\pi}{p^{*2}} \sum_{l=0}^{\infty} (2l+1)(1 - |\eta_l|^2), \tag{2.22}$$

where $|\eta_l|$ is the amplitude of the l^{th} wave in the expansion, for which $0 \leq |\eta_l| \leq 1$ holds (unitarity). Because in the Fermi theory interactions are point-like, only the s-wave has to be considered ($l = 0$). This leads to the following constraint for the scattering cross section:

$$\sigma \leq \frac{\pi}{p^{*2}} \sum_{l=0}^{\infty} (2l+1) \tag{2.23}$$

$$\Rightarrow \sigma \leq \frac{\pi}{p^{*2}}. \tag{2.24}$$

A comparison to the cross section of the Fermi interaction results in a boundary for the

center-of-mass energy and the particle momenta:

$$\frac{G_F^2}{\pi} \cdot 4p^{*2} \leq \frac{\pi}{p^{*2}} \tag{2.25}$$

$$p^* \leq \sqrt{\frac{\pi}{2G_F}} \sim 370\,\mathrm{GeV}, \tag{2.26}$$

with $G_F = 1.166 \cdot 10^{-5}\,\mathrm{GeV}^{-2}$. As a consequence, the Fermi theory has to be replaced by a new model for weak interactions at energies $\gtrsim 370\,\mathrm{GeV}$, which should preferentially be – in contrast to the Fermi theory – a renormalizable gauge theory [2, 3, 17].

Electroweak Interactions

The theory of electroweak interactions was developed in the 1950s and 1960s with the efforts of S. Glashow [18], A. Salam [19] and S. Weinberg [20], who were awarded with the Nobel Prize in 1979.

Weak interactions can transform e.g. left-handed electrons into left-handed neutrinos and left-handed u-quarks into left-handed d'-quarks as if they were two states of the same particle. All of them carry weak isospin $I = \frac{1}{2}$ and thus form weak isospin doublets

$$L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \begin{pmatrix} u_L \\ d'_L. \end{pmatrix} \tag{2.27}$$

Right-handed fermions have weak isospin $I = 0$ and are singlets under the weak interaction:

$$\left(e\right)_R, \left(u\right)_R, \left(d'\right)_R. \tag{2.28}$$

This means that weak interactions can be understood as rotations in isospin space which only affect the left-handed doublets:

$$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}' = U \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{2.29}$$

$$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}' = e^{\frac{i}{2}\vec{\alpha}\vec{\tau}} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \tag{2.30}$$

where $\vec{\alpha}$ denotes the rotation angles with respect to the three weak isospin axes and $\vec{\tau}$ is a column vector containing the three Pauli matrices. The symmetry we thus need local gauge invariance for is that of $SU(2)$. In order to take into account parity violation, one can introduce an additional chirality operator $P_L = \frac{1}{2}(1-\gamma^5)$ to those rotations, ensuring that only left-handed fermions participate in the interaction. Therefore, the symmetry group is called $SU(2)_L$.

However, it turns out that two separate gauge theories for weak and electromagnetic interactions like $SU(2)_L \otimes U(1)_{\mathrm{QED}}$ cannot be formulated as the commutation of both operations leads to inconsistencies, which would require the charges of e.g. the neutrino and the electron to be the same. This is certainly not the case in the electromagnetic

force. A new $U(1)$ symmetry has to be created such that the electron and neutrino yield the same charge. Therefore the weak hypercharge is introduced as

$$Y = 2(Q - I_3). \tag{2.31}$$

Electrons and neutrinos indeed have identical weak hypercharge. One can now continue to construct a gauge theory based on the $U(1)_Y$ symmetry in the same way it was performed for QED. The resulting gauge boson is commonly called B-boson. It is electrically neutral and couples with coupling strength g' and in dependence on the weak hypercharge. This also means that the B-boson couples with different strength to left-handed and right-handed particles.

For gauging the Dirac equation for local $SU(2)$ transformations in the isospin space, the Dirac spinors have to be supplemented with an isospin part:

$$\psi = \left[\omega_\nu \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \omega_e \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] \psi_{\text{Dirac}}. \tag{2.32}$$

In order to compensate rotations of the form of Eqn. 2.30, three gauge fields are needed:

$$\vec{W}_\mu(x) = \begin{pmatrix} W^1_\mu(x) \\ W^2_\mu(x) \\ W^3_\mu(x) \end{pmatrix}, \tag{2.33}$$

which enter the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + \frac{ig}{2}\vec{W}_\mu(x)\vec{\tau}, \tag{2.34}$$

with g being the coupling constant. The $\vec{W}_\mu(x)$ fields obey the following transformation rule:

$$\vec{W}_\mu(x) \to \vec{W}'_\mu(x) = \vec{W}_\mu(x) + \frac{1}{g}\partial_\mu\vec{\alpha}(x) - \vec{W}_\mu(x) \times \vec{\alpha}(x). \tag{2.35}$$

Here it has to be noted that the cross product of the $\vec{W}_\mu$ fields with the vector of rotation angles $\vec{\alpha}$ will give rise to self-coupling terms of the W-bosons in the Lagrangian. The electroweak Lagrangian then reads so far:

$$\mathcal{L}_{SU(2)_L \otimes U(1)_Y} = \underbrace{\bar{L} i\gamma^\mu \partial_\mu L + \bar{e_R} i\gamma^\mu \partial_\mu e_R}_{\text{free fermion part}} \tag{2.36}$$

$$\underbrace{-\bar{L}\left(\frac{1}{2}g' Y_L \gamma^\mu B_\mu\right)L - \bar{e_R}\left(\frac{1}{2}g' Y_R \gamma^\mu B_\mu\right)e_R}_{\text{fermion interaction with } B_\mu} \tag{2.37}$$

$$\underbrace{-\bar{L}\left(\frac{1}{2}g\gamma^\mu \vec{\tau}\vec{W}_\mu\right)L}_{\text{fermion interaction with } \vec{W}_\mu} \underbrace{-\frac{1}{4}B_{\mu\nu}B^{\mu\nu}}_{\text{free } B_\mu} \underbrace{-\frac{1}{4}\vec{W_{\mu\nu}}\vec{W^{\mu\nu}}}_{\text{free } \vec{W}_\mu}, \tag{2.38}$$

where $B_{\mu\nu}$ is defined in analogy to $F_{\mu\nu}$ in QED and $\vec{W_{\mu\nu}}$ is given as

$$\vec{W_{\mu\nu}} = \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu - g\vec{W}_\mu \times \vec{W}_\nu, \tag{2.39}$$

of which the last term includes self-coupling diagrams of the $\vec{W}_\mu$-fields.

Two of the W-bosons can be identified as the charged currents $W^\pm$:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2). \tag{2.40}$$

However, the short interaction range of the weak force and experimental data show that the W-bosons have to carry mass which has been neglected so far in this discussion. Furthermore, the B- and W_μ^3-bosons have not yet been set into relation with the experimentally found photon and Z^0-boson. The solution for these issues comes with the introduction of the Higgs mechanism [2, 9, 12].

Higgs Mechanism

What is known today as the Higgs mechanism was originally developed by several scientists in parallel: F. Englert and R. Brout [21], P. Higgs [22] and G. Guralnik, C. Hagen and T. Kibble [23]. The discovery of the massive boson it predicted at the ATLAS [24] and CMS experiments [25] LHC in 2012 lead to a Nobel Prize in 2013.

Consider a complex scalar field $\Phi(x) = \frac{1}{\sqrt{2}}\big(\phi_1(x) + i\phi_2(x)\big)$ (with real ϕ_1 and ϕ_2) and a potential $V(\Phi)$ which is invariant under rotations in the plane spanned by ϕ_1 and ϕ_2:

$$V(\Phi) = -\mu^2|\Phi|^2 + \lambda^2|\Phi|^4. \tag{2.41}$$

While the potential itself is invariant under $U(1)$ transformations the ground state $|\Phi_0| = \frac{v}{\sqrt{2}} = \frac{\mu}{\sqrt{2}\lambda}$ is not – a principle which is known as spontaneous symmetry breaking. The Lagrangian for this system then reads:

$$\mathcal{L} = (\partial_\mu\Phi)(\partial_\mu\Phi)^* - V(\Phi). \tag{2.42}$$

$\Phi(x)$ can be parameterized by small oscillations around the ground state:

$$\Phi(x) = \frac{1}{\sqrt{2}}\big(v + \eta(x) + i\zeta(x)\big), \tag{2.43}$$

where $\eta(x)$ represents oscillations in radial direction and $\zeta(x)$ in azimuthal direction. Entering this representation of Φ into the potential, the Lagrangian yields:

$$\mathcal{L} = \left(\frac{1}{2}\big(\partial_\mu\eta(x)\big)\big(\partial^\mu\eta(x)\big) - \mu^2\eta^2\right) + \left(\frac{1}{2}\big(\partial_\mu\zeta(x)\big)\big(\partial^\mu\zeta(x)\big)\right) + \text{higher orders in } \eta \text{ and } \zeta. \tag{2.44}$$

This can be interpreted as the Lagrangians for a massive scalar boson $\eta(x)$ with $m = \sqrt{2}\mu$ and a massless (Nambu-Goldstone[1]) boson $\zeta(x)$.

The same scheme can be applied to the electroweak theory to give a theoretical foundation for the vector boson masses. For that purpose, the simplest way is to introduce the field Φ as an isospin doublet with four scalar fields:

$$\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}, \tag{2.45}$$

[1] According to the Goldstone theorem, a massless scalar particle arises if a continuous symmetry is spontaneously broken [26–28]. This particle is referred to as Nambu-Goldstone boson.

with weak isospin $I = \frac{1}{2}$ and weak hypercharge $Y = 1$. The important parts in the Lagrangian are now the inserted potential $V(\Phi^\dagger\Phi)$, which is defined in analogy to Eqn. 2.41, and the kinetic part $(D_\mu\Phi)^\dagger(D^\mu\Phi)$ with the covariant derivative of $SU(2)_L \otimes U(1)_Y$:

$$D_\mu = \partial_\mu + \frac{1}{2}ig'YB_\mu(x) + \frac{1}{2}ig\vec{\tau}\vec{W}_\mu(x). \tag{2.46}$$

The ground state of $V(\Phi^\dagger\Phi)$ can be chosen as

$$\Phi_0 = \begin{pmatrix}\phi_0^+\\ \phi_0^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{2.47}$$

with a real constant parameter v. Again, $\Phi(x)$ can be written as an expansion around the ground state. Inserting this expansion into the potential and kinetic terms of the Lagrangian and identifying the photon and Z-boson as

$$A_\mu(x) = B_\mu(x)\cos\theta_W + W_\mu^3(x)\sin\theta_W \tag{2.48}$$

$$Z_\mu(x) = -B_\mu(x)\sin\theta_W + W_\mu^3(x)\cos\theta_W, \tag{2.49}$$

respectively, with the Weinberg angle θ_W, the Lagrangian takes the following form:

$$\mathcal{L}_{\text{Higgs}} = \underbrace{\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2}_{\text{free massive Higgs boson}} + \underbrace{\frac{1}{2}\frac{g^2v^2}{4\cos^2(\theta_W)}|Z_\mu|^2}_{Z\text{-boson mass term}} + \underbrace{\frac{1}{2}\frac{g^2v^2}{4}\left(|W_\mu^+|^2 + |W_\mu^-|^2\right)}_{W\text{-boson mass term}} \tag{2.50}$$

$$+ \text{Higgs self coupling terms} + \text{Interaction terms of Higgs, } W_\mu \text{ and } Z_\mu. \tag{2.51}$$

Here, three of the initially four new degrees of freedom have been absorbed by the electroweak gauge fields through exploiting gauge invariance, thereby giving them mass and a third polarization state. The fourth degree of freedom gives rise to a massive scalar – the so-called Higgs boson. From Eqn. 2.50, the masses of the $W^\pm$-, Z^0- and Higgs boson masses can be written as [2, 9, 12]:

$$m_W = \frac{gv}{2} \tag{2.52}$$

$$m_Z = \frac{gv}{2\cos\theta_W} = \frac{m_W}{\cos\theta_W} \tag{2.53}$$

$$m_H = \sqrt{2}\mu \tag{2.54}$$

In addition to giving an explanation for the origin of the vector boson masses in the electroweak sector, the Higgs field is also the cause for the fermion masses in the Standard Model by what is known as Yukawa coupling.

Considering the Dirac equation for an electron and applying the chirality operator P_R, it can be observed that the left-handed chirality component is related to the right-handed component:

$$P_R(i\gamma^\mu\partial_\mu - m)\psi_e = 0 \tag{2.55}$$

$$i\gamma^\mu\partial_\mu \underbrace{\frac{1}{2}(1-\gamma^5)\psi_e}_{P_L\psi_e \equiv e_L} - m\underbrace{\frac{1}{2}(1+\gamma^5)\psi_e}_{P_R\psi_e \equiv e_R} = 0 \tag{2.56}$$

$$i\gamma^\mu\partial_\mu e_L = m \cdot e_R \tag{2.57}$$

The mass term for the electron should have a form similar to the expression in Eqn. 2.52. Taking into account $SU(2)_L$ symmetry as well, this can also be written in terms of isospin doublets and the coupling constant $\tilde{g}_e$ to the Higgs field:

$$i\gamma^\mu \partial_\mu \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L = \frac{\tilde{g}_e}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} e_R, \tag{2.58}$$

where the ground state of the Higgs field $\Phi(x)$ has been used. The neutrino mass is then 0 (which was falsified by the discovery of neutrino oscillations [29, 30]) and the electron's mass is proportional to the coupling constant $\tilde{g}_e$. Such coupling constants have to be introduced for every massive fermion in the Standard Model as free parameters to the theory. Using the shorthand of L for the isospin doublet and R for the singlet, the corresponding addition to the Standard Model Lagrangian is the following term

$$\mathcal{L}_{\text{Yukawa}} = -\tilde{g}_f \big(\bar{R}(\Phi^\dagger L) + (\bar{L}\Phi)R\big), \tag{2.59}$$

for each fermion with mass $\frac{\tilde{g}_f v}{\sqrt{2}}$ [2, 17].

Quantum Chromodynamics

The theory of strong interactions is a gauge theory that is based on the invariance under local $SU(3)$ transformations. Quarks carry charge with respect to strong interaction, which can appear in three different colours (red, green and blue) or anti-colours (anti-red, anti-green and anti-blue) for anti-quarks. To represent this, the Dirac spinors for quarks are complemented by a colour vector:

$$\Psi = \underbrace{\psi(x)}_{\text{Dirac spinor}} \cdot \underbrace{\chi_{\text{colour}}}_{\text{colour vector}} \tag{2.60}$$

$$\chi_{\text{colour}} = \alpha\chi_R + \beta\chi_G + \gamma\chi_B \tag{2.61}$$

$$\chi_R = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \chi_G = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \chi_B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.62}$$

Consider a $SU(3)$ local gauge transformation

$$\Psi(x) \to \Psi'(x) = e^{-\vec{\alpha}(x)\frac{\vec{\lambda}}{2}}\Psi(x) \tag{2.63}$$

with $\vec{\lambda}$ being the 8 $SU(3)$ generators, which can be represented by the Gell-Mann matrices. To compensate for this transformation eight gauge fields have to be introduced, namely the gluon fields $\vec{G}_\mu = (G^1_\mu, ..., G^8_\mu)$, in the covariant derivative:

$$\partial_\mu \to D_\mu = \partial_\mu + \frac{1}{2} i g_s \vec{\lambda}\vec{G}_\mu(x), \tag{2.64}$$

with the strong coupling constant g_s. The gluon gauge fields transform as

$$\vec{G}_\mu(x) \to \vec{G}'_\mu(x) = \vec{G}_\mu(x) + \frac{1}{g_s}\partial_\mu\vec{\alpha}(x) - \vec{G}_\mu(x) \times \vec{\alpha}(x). \tag{2.65}$$

In analogy to the electroweak interaction theory, the field strength tensor for the gluon fields is defined as the following:

$$\vec{G}^{\mu\nu} = \partial^\nu\vec{G}^\mu - \partial^\mu\vec{G}^\nu - g_s\vec{G}^\mu \times \vec{G}^\nu, \tag{2.66}$$

where the last term gives rise to gluon self coupling diagrams. The QCD Lagrangian thus takes the following form:

$$\mathcal{L} = \underbrace{\bar{\Psi}\left(i\gamma_\mu D^\mu - m\right)\Psi}_{\text{free quark and quark-gluon interaction}} \underbrace{-\frac{1}{4}\vec{G}_{\mu\nu}\vec{G}^{\mu\nu}}_{\text{free gluons}} \tag{2.67}$$

In contrast to weak interactions, the gauge bosons of strong interactions remain massless. Furthermore, QCD shows some peculiarities in the behaviour of its coupling constant depending on the transferred momentum Q:

$$\alpha_s(Q^2) = \frac{g_s^2}{4\pi} = \frac{12\pi}{(33 - 2N_f)\ln\left(\frac{Q^2}{\Lambda^2}\right)}(1 + ...), \tag{2.68}$$

with N_f the number of quark flavours, that can kinematically participate, and a free parameter Λ serving as a cut-off scale, below which the perturbative evolution of QCD breaks down. This dependence of the coupling constant on the transferred momentum is not that surprising as it arises from vacuum polarization diagrams, for which the emerging infinities are absorbed in renormalized effective charges and couplings, similar to what was discussed in the section for QED. However, looking at g_s for very high and low momentum transfers, one can observe the following:

$$\alpha_s(Q^2) \xrightarrow{Q^2 \to \infty} 0 \tag{2.69}$$

$$\alpha_s(Q^2) \xrightarrow{Q^2 \to \Lambda^2} \infty. \tag{2.70}$$

Eqn 2.69 is usually interpreted as quarks behaving similarly to free particles at very high energies (asymptotic freedom). In contrast, Eqn. 2.70 hints that the perturbative nature of QCD is not valid anymore at small energy scales and that all strongly interacting particles are bound in colour neutral systems (hadrons), which is referred to as confinement [2, 9, 10, 17].

2.2. Shortcomings of the Standard Model of Particle Physics

The Standard Model of particle physics is so far the best available description of subnucleonic processes. However, it has some intrinsic shortcomings, e.g. there is no motivation for the hierarchy of particle masses as well as there is no explanation for why there should be exactly three generations of quarks and leptons. Furthermore, to date it was not possible to unify the $SU(2)_L \otimes U(1)_Y$ theory of electroweak interactions with the $SU(3)_C$ gauge theory of strong interaction. Gravity has not been incorporated into the theory in any way so far. There are also many observations in astrophysics and cosmology for which the Standard Model of particle physics currently does not deliver any explanation. Because of these reasons, an extension or modification of the Standard Model will be necessary. The common belief is that this new theory will come into play at some higher energies with the current Standard Model being a low-energy approximation of this more fundamental theory. Some of the candidates for such a theory rely on Supersymmetry, which will be discussed in the next chapter after first highlighting a few of the major shortcomings of the Standard Model, for which there are solutions in supersymmetric models.

2.2.1. Unification of Interactions

In the Standard Model, the different interactions are simply pieced together into one framework. In the light of the running of coupling constants, however, this is probably not the most fundamental description. For high energies the renormalization group equations predict that the values of the coupling constants start to converge. The solution of the renormalization group equations is given as

$$\frac{1}{\tilde{\alpha}_i(Q^2)} = \frac{1}{\tilde{\alpha}_i(\mu^2)} - b_i \ln \frac{Q^2}{\mu^2}, \tag{2.71}$$

with the index $i = 1, 2, 3$ standing for electromagnetic interaction (1), weak interaction (2) and strong interaction (3), respectively. μ^2 is one particular value in the Q^2 range, at which the couplings have been measured. The coupling constants $\tilde{\alpha}_i$ are defined as

$$\tilde{\alpha}_i = \frac{\alpha_i}{4\pi} = \frac{g_i^2}{(4\pi)^2}. \tag{2.72}$$

The important factors are the coefficients b_i, which for the Standard Model are given as

$$b_i = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -22/3 \\ -11 \end{pmatrix} + N_{\text{generations}} \begin{pmatrix} 4/3 \\ 4/3 \\ 4/3 \end{pmatrix} + N_{\text{Higgs}} \begin{pmatrix} 1/10 \\ 1/6 \\ 0 \end{pmatrix}, \tag{2.73}$$

with $N_{\text{generations}}$ the number of quark and lepton generations (3) and N_{Higgs} the number of Higgs doublets in the Standard Model (1). The problem is, with the ingredients of the Standard Model the couplings come very close together at high Q^2 but miss each other, as can be observed in the left plot of Fig. 2.1. There are, however, theories which predict a common intersection point of all the coupling constants. In a minimal supersymmetric model as outlined in Sec. 2.3.3 the coefficients b_i are different:

$$b_i = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \\ -9 \end{pmatrix} + N_{\text{generations}} \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + N_{\text{Higgs}} \begin{pmatrix} 3/10 \\ 1/2 \\ 0 \end{pmatrix}, \tag{2.74}$$

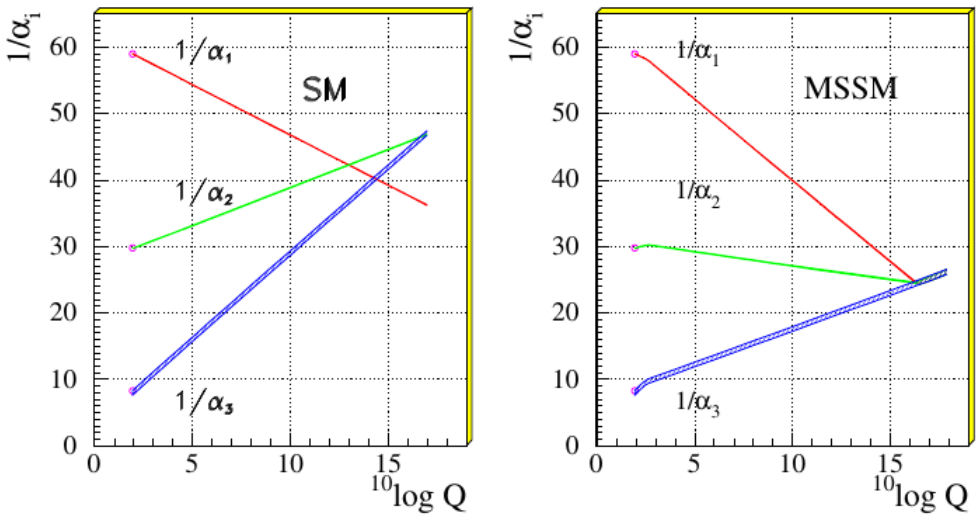


Figure 2.1.: These plots depict the running of coupling constants for the Standard Model (left) and the minimal supersymmetric extension of the Standard Model (right). The left plot is based on measurements of α_i at $Q = M_Z$. For the right plot this has been extrapolated to $Q = M_{\mathrm{SUSY}}$, where M_{SUSY} has been varied in a fit. SUSY particles only contribute above M_{SUSY} which causes the change in the slope of the couplings in the right plot. [31, 32].

with the number of quark and lepton generations again being 3 and 2 Higgs isospin doublets. With this set of parameters, the coupling constants now merge in a single value at an energy scale of $M_{\mathrm{GUT}} = 10^{15}$ to $10^{16}\,\mathrm{GeV}$, which makes a unified description of all interactions possible. This is shown in the right plot of Fig. 2.1. The minimal "grand unified theory" (GUT) is a $SU(5)$ gauge group, but there are also other candidates as e.g. $SO(10)$ [2, 10, 31, 32].

Another obvious deficit of the Standard Model is the complete neglect of the gravitational force, as it is much weaker compared to the other known interactions and thus only comes into play at extremely high energy scales, i.e. the Planck scale ($10^{19}\,\mathrm{GeV}$). The attempts to create a quantum field theory for gravity have had one major fault so far: It is non-renormalizable. Furthermore, the Coleman-Mandula theorem forbids a non-trivial unification of the Lorentz group with a compact internal symmetry group. The unification with the other forces in the Standard Model is therefore not possible. Supersymmetry could also be the key to solve this puzzle: It will be seen in Sec. 2.3.2 that Supersymmetry provides exactly such a link between the Lorentz group and internal symmetries. Furthermore, if Supersymmetry is taken as a local gauge symmetry, it also becomes a theory of gravity – so-called supergravity [2, 10].

2.2.2. Hierarchy Problem

The hierarchy problem is a challenge that is not so much faced by the Standard Model itself but rather by the expectation that there has to be some new physics at higher mass scales, like the GUT scale ($10^{15}\,\mathrm{GeV}$) and ultimately also Planck scale ($10^{19}\,\mathrm{GeV}$). Consider an $SU(5)$ GUT model as an example, which has two Higgs bosons with masses M_H somewhere

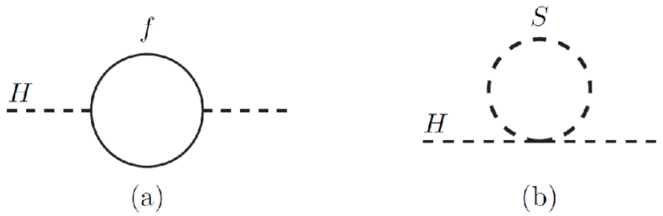


Figure 2.2.: Depicted here are the quantum loop corrections to the parameter μ^2 caused by a Dirac fermion (a) and a scalar boson (b). Because they enter with different signs, their effects on the Higgs mass cancel. [33]

at the electroweak scale of $10^2\,\mathrm{GeV}$ and M_X located at the GUT scale: The difference between those two mass scales will be disturbed by the loop corrections that will add to the lighter Higgs mass and thus decrease the energy gap. In principle, such one-loop corrections could be reduced by "tuning" the parameters of the Higgs potential, in order to restore the mass scale hierarchy. Yet, there will be a scale mixing introduced by two-loop diagrams instead. Ironing out the mass scale mixing by fine-tuning the parameters at each order of perturbation theory appears, however, not natural. Again, Supersymmetry provides an "automatic" solution to the hierarchy problem: By introducing fermionic and bosonic states in the same number, and with fermionic and bosonic loop corrections having opposite sign, the corrections will cancel out each other without the need for extensive fine-tuning. An illustration of this mechanism is given in Fig. 2.2 [2, 10, 33].

2.2.3. Dark Matter

There are several cosmological and astrophysical observations proving that our knowledge only applies to a small fraction of the universe's content. The particles described by the Standard Model only account for approximately 5% of the universe's content, the remaining fraction is predominantly due to dark energy (69%) and dark matter (26%). While the nature of dark energy is far from being understood, there are experimental indications and theories that give at least a hint to what dark matter could be composed of. Dark matter is known to interact via the gravitational force, which is also the way it manifests itself in astrophysical measurements and observations. Most galaxies' rotational behaviour, for example, cannot be solely explained by ordinary matter: The angular velocities especially of the stars at the outskirts of galaxies are far too high and actually should tear them apart. Similar observations have been made for entire galaxies in galaxy clusters. Another example would be the observation of gravitational lensing in the absence of any sufficient amounts of luminous matter. All of this indicates that there is another source of gravity, a kind of matter that does not interact electromagnetically and that telescopes therefore cannot detect - so-called dark matter. It would, however, be possible that dark matter particles interact via the weak force. In the case of "cold" dark matter, these particles are also expected to have considerable mass $\gtrsim 30\,\mathrm{keV}$. A general term for a hypothetical particle like this is "weakly interacting massive particle" or for short WIMP.

Early in the universe in a regime of very high temperatures and densities, these WIMPs would exist in thermal equilibrium with ordinary matter. The expansion and cooling of the universe, however, would lead to a thermal freeze-out of the dark matter particles, i.e.

no new WIMPs could be produced. Annihilation processes of the WIMPs also decreased due to further cosmic expansion and the lowering of the particle density that came with it. Considering the WIMP scenario, in order to obtain the dark matter density inferred from experiment, the effective annihilation cross section times the relative speed of the WIMPs should be

$$\langle \sigma v \rangle \sim 1\,\mathrm{pb} \sim \frac{\alpha^2}{(150\,\mathrm{GeV})^2}, \tag{2.75}$$

which is roughly consistent with the weak mass scale, with α being the weak coupling constant.

There are several possibilities for the theoretical origin of the WIMPs, some of them are supersymmetric. As dark matter particles should be stable, the lightest supersymmetric particle would be an obvious candidate for a WIMP, given an R-parity conserving Supersymmetry scenario (cf. Sec. 2.3.4) and a neutral lightest supersymmetric particle, e.g. the lightest neutralino. The lightest neutralino arises through mixing of the neutral Higgsinos, winos and binos. If the lightest neutralino is indeed the particle that dark matter is mainly composed of, conclusions can be drawn concerning the mixing. If the mixing is dominated by the Higgsinos or winos (i.e. the neutralino is Higgsino- or wino-like) and has a mass below 1 TeV, the annihilation and co-annihilation mechanisms via a sfermion are expected to be much too efficient to match the observed dark matter density. Although the annihilation cross section for bino-like neutralinos would probably be too small, there are some scenarios that would reconcile this fact with observed the dark matter density [33, 34].

2.3. Supersymmetry

2.3.1. Original Motivation for Supersymmetry

In 1967, Coleman and Mandula published a paper, in which they derived that it was impossible to combine internal and space-time symmetries in a non-trivial way ("no-go theorem") [35]. However, during the very same year, Hironari Miyazawa also made a publication on a new symmetry, in which baryons and mesons are forming a supermultiplet – the first mentioning of a supersymmetric group [36], with which exactly was achieved what Coleman and Mandula forbid: A combination of internal and non-compact space-time groups. In the 1970s, quite some progress was made in the field of supersymmetric theories, however, its value was for a long time considered as a purely mathematical one. When attempts were made to get rid of the inconsistencies of GUT theories, it was realized that Supersymmetry could be the key. Another advantage of Supersymmetry is that it yields theories which are finite to all orders of perturbation theory, and thus without the need of renormalization [10].

2.3.2. Supersymmetric Algebra

Supersymmetry relates bosonic ($|B\rangle$) to fermionic states ($|F\rangle$) and vice versa. To achieve this, a spin $\frac{1}{2}$ operator Q is introduced, such that

$$Q|B\rangle = |F\rangle. \tag{2.76}$$

Consider the Hamiltonian H with the bosonic and fermionic operators a and b

$$H = \omega_a a^\dagger a + \omega_b b^\dagger b, \tag{2.77}$$

and the operators obeying the (anti-)commutation relations

$$\left[a, a^\dagger\right] = \left\{b, b^\dagger\right\} = 1. \tag{2.78}$$

Then, the operator for supersymmetric transformation Q can be defined as

$$Q \equiv b^\dagger a + a^\dagger b, \tag{2.79}$$

such that Q has the same effect as was sketched qualitatively in Eqn. 2.76. Q relates in the following way to the Hamiltonian:

$$\left[Q, H\right] = (\omega_a - \omega_b)Q, \tag{2.80}$$

i.e. the supersymmetric operator and the Hamiltonian commute if the fermions and bosons have equal energy ($\omega_a = \omega_b = \omega$) and

$$\left\{Q, Q^\dagger\right\} = \frac{2}{\omega}H \tag{2.81}$$

holds. An interesting observation is that in this case the energy of the supersymmetric vacuum must be 0. This is different in broken supersymmetric theories, however, as there the vacuum will yield positive energy.

One of the simplest supersymmetric Lagrangians that can be constructed, is the one found by Gervais and Sakita in 1971, which contains the kinetic terms of a fermionic Weyl-spinor ψ and a complex scalar field ϕ forming a supermultiplet:

$$\mathcal{L} = \bar{\psi}^a(i\gamma^\mu\partial_\mu)\psi^a + \partial_\mu\phi^a\partial^\mu\phi^a. \tag{2.82}$$

The Lagrangian is invariant under the supersymmetric transformation

$$\delta\psi^a = -i\gamma^\mu\partial_\mu\phi^a\epsilon \tag{2.83}$$

$$\delta\phi^a = \bar{\epsilon}\psi^a, \tag{2.84}$$

where the ϵ is an anti-commuting spinor, which parameterizes the transformation. If this transformation is carried out twice, one finds the following result

$$\delta_1(\delta_2\phi^a) = \bar{\epsilon_1}\gamma^\mu\epsilon_2(-i\partial_\mu)\phi^a \tag{2.85}$$

$$= \bar{\epsilon_1}\gamma^\mu\epsilon_2 P_\mu\phi^a. \tag{2.86}$$

Here, $-i\partial_\mu$ has been identified with the energy-momentum operator P_μ. This can be interpreted as the "square" of the supersymmetric operator generating a Lorentz-boost – a connection which is often characterized by the following anti-commutation relation:

$$\left\{Q_\alpha, Q^\dagger_\beta\right\} = -2\sigma^\mu_{\alpha\beta}P_\mu, \tag{2.87}$$

where Q_α with $\alpha = 1, 2$ are the left-handed spinor components of the supersymmetric operator and $Q^\dagger_\beta$ are the right-handed components [10, 11, 33].

2.3.3. Minimal Supersymmetric Extension to the Standard Model (MSSM)

In a minimal approach to extend the Standard Model in a supersymmetric way, the known fermions and bosons have to be arranged into supermultiplets, together with their newly introduced superpartners. Each supermultiplet contains an equal number of fermionic and bosonic states. There are several different ways to construct such a supermultiplet. One of them are the so-called gauge supermultiplets, containing a vector boson, carrying spin 1 and no mass (to keep the theory renormalizable at first, but masses can then be obtained by spontaneous symmetry breaking). A massless vector boson has two polarization states, therefore there are two degrees of freedom. The superpartner must match the number of degrees of freedom and is therefore a massless spin $\frac{1}{2}$ Weyl fermion with two helicity states. If the vector boson is one of the Standard Model gauge bosons (e.g. the gluon) the fermionic partner particle in the supermultiplet is called a gaugino (e.g. gluino). These fermions must obey the same transformation rules as their gauge boson superpartners, irrespective of their chirality state (in contrast to the Standard Model fermions).

If gravity is included, the graviton, a spin 2 boson, would obtain a spin $\frac{3}{2}$ partner, the gravitino.

The Standard Model fermions are placed into chiral supermultiplets. Keeping in mind that the quarks and leptons in the Standard Model behave differently under the electroweak gauge transformations, their superpartners, although being spin 0 scalars, inherit this behaviour. For instance, the left-handed leptons' superpartners (sleptons) will couple to the W-boson, while the superpartners of the right-handed leptons won't, although they do not have the property of helicity themselves. Hence, the nomenclature "honours" this

fact by adding an index "L" ("R") to the superpartners of left-handed (right-handed) fermions.

The Higgs isospin doublet, as it is a scalar particle itself, should be part of a chiral supermultiplet as well and have a fermionic partner (Higgsino). It turns out that, in order to avoid gauge anomalies, one Higgs isospin doublet is not enough. Two such complex doublets are needed: One $Y = \frac{1}{2}$ doublet, $H_u = (H_u^+, H_u^0)$, which creates Yukawa couplings for the up-type quarks with $+\frac{2}{3}$ electric charge and a $Y = -\frac{1}{2}$ doublet, $H_d = (H_d^0, H_d^-)$, generating Yukawa couplings for the down-type quarks with electric charge $-\frac{1}{3}$ and for the charged leptons. The Standard Model Higgs boson h^0 corresponds to a linear combination of H_u^0 and H_d^0. In electroweak symmetry breaking, three of the eight degrees of freedom are Goldstone bosons, that are absorbed by the Z- and W- bosons. The other five form the CP-even neutral scalar bosons h^0 and H^0, the CP-odd neutral scalar boson A^0 and two charged scalars $H^\pm$. The chiral and gauge supermultiplets of the MSSM are listed in Tab. 2.2.

Electroweak symmetry breaking is also the cause for the mixing of the neutral Higgsino states and the neutral electroweak gauginos $\tilde{B}^0$ and $\tilde{W}_3$ to form in total four neutral fermions, called neutralinos. They are ordered according to their masses from $\tilde{\chi}_1^0$ (lightest) to $\tilde{\chi}_4^0$ (heaviest). The lightest neutralino is very often supposed to be the lightest supersymmetric particle and therefore poses a good candidate for dark matter. Similarly, there are four mixed mass eigenstates of the charged Higgsinos and the charged winos, which are known as charginos $\tilde{\chi}_{1,2}^\pm$.

Mixing also occurs for squarks and sleptons. The mixing between generations is in general assumed to be very small as the mixing in the first two generations is believed to be in general, while the partners of the left-handed and right-handed Standard Model fermions especially in the third generation can in principle show substantial mixing. For example, the scalar superpartners of the left- and right-handed τ-lepton, $\tilde{\tau}_L$ and $\tilde{\tau}_R$, can mix to give the mass eigenstates

$$\begin{pmatrix} \tilde{\tau}_1 \\ \tilde{\tau}_2 \end{pmatrix} = \begin{pmatrix} \cos\theta_{\tilde{\tau}} & -\sin\theta_{\tilde{\tau}} \\ \sin\theta_{\tilde{\tau}} & \cos\theta_{\tilde{\tau}} \end{pmatrix} \begin{pmatrix} \tilde{\tau}_L \\ \tilde{\tau}_R \end{pmatrix}, \tag{2.88}$$

assuming real off-diagonal elements in the mixing matrix and using the mixing angle $\theta_{\tilde{\tau}}$, which can be in the range $0 \leq \theta_{\tilde{\tau}} \leq \pi$.

The additional particle content of the MSSM (not including the already known Standard Model particles) in terms of mass eigenstates is given in Tab. 2.3.

The MSSM introduces 105 new free parameters, which have to be determined by experiment. This is of course highly difficult and impractical when searching for supersymmetric processes in experiments. Therefore, a "reduced" version of the MSSM was developed, the phenomenological minimal extension to the Standard Model (pMSSM), which includes a total of 19 free parameters. However, this is still a very large parameter space. There a several strategies in searching for Supersymmetry. One of them is to scan the pMSSM (or any other fully-fledged Supersymmetry model) and evaluate the data in each set of chosen parameters. Another approach is to focus only on one or a few production diagrams of supersymmetric particles, assuming that other supersymmetric processes are suppressed and other supersymmetric particles than the ones considered are kinematically decoupled. Therefore a 100% branching fraction for the chosen process is assumed. The scan is then only performed in a phase-space spanned only by very few parameters – typically the masses or mass differences of the sparticles involved. The results of the analyses of such

"simplified models" can then serve as input or constraints for the development of new supersymmetric models [2, 33].

Names		spin 0	spin 1/2
squarks, quarks	Q	$(\tilde{u}_L,\, \tilde{d}_L)$	$(u_L,\, d_L)$
	$\bar{u}$	$\tilde{u}_{R}{}^*$	$u_R^\dagger$
	$\bar{d}$	$\tilde{d}_{R}{}^*$	$d_R^\dagger$
sleptons, leptons	L	$(\tilde{\nu},\, \tilde{e}_L)$	$(\nu,\, e_L)$
	$\bar{e}$	$\tilde{e}_{R}{}^*$	$e_R^\dagger$
Higgs, Higgsinos	H_u	$(H_u^+,\, H_u^0)$	$(\tilde{H}_u^+,\, \tilde{H}_u^0)$
	H_d	$(H_d^0,\, H_d^-)$	$(\tilde{H}_d^0,\, \tilde{H}_d^-)$

Names	spin 1/2	spin 1
gluino, gluon	$\tilde{g}$	g
winos, W bosons	$\tilde{W}^\pm,\, \tilde{W}^0$	$W^\pm,\, W^0$
bino, B boson	$\tilde{B}^0$	B^0

Table 2.2.: The chiral and gauge supermultiplets in the minimal supersymmetric extension of the Standard Model (MSSM). It is important to keep in mind that the helicity indices (R, L) of the squarks and sleptons are simply inherited from their Standard Model partners and do not refer to any helicity state as they are spin-0 particles [33].

Name	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$H_u^0,\, H_d^0,\, H_u^+,\, H_d^-$	$h^0,\, H^0,\, A^0,\, H^\pm$
Squarks	0	-1	$\tilde{u}_L,\, \tilde{u}_R,\, \tilde{d}_L.\ \tilde{d}_R$ $\tilde{c}_L,\, \tilde{c}_R,\, \tilde{s}_L,\, \tilde{s}_R$ $\tilde{t}_L,\, \tilde{t}_R,\, \tilde{b}_L,\, \tilde{b}_R$	$\tilde{u}_L,\, \tilde{u}_R,\, \tilde{d}_L.\ \tilde{d}_R$ $\tilde{c}_L,\, \tilde{c}_R,\, \tilde{s}_L,\, \tilde{s}_R$ $\tilde{t}_1,\, \tilde{t}_2,\, \tilde{b}_1,\, \tilde{b}_2$
Sleptons	0	-1	$\tilde{e}_L,\, \tilde{e}_R,\, \tilde{\nu}_e$ $\tilde{\mu}_L,\, \tilde{\mu}_R,\, \tilde{\nu}_\mu$ $\tilde{\tau}_L,\, \tilde{\tau}_R,\, \tilde{\nu}_\tau$	$\tilde{e}_L,\, \tilde{e}_R,\, \tilde{\nu}_e$ $\tilde{\mu}_L,\, \tilde{\mu}_R,\, \tilde{\nu}_\mu$ $\tilde{\tau}_1,\, \tilde{\tau}_2,\, \tilde{\nu}_\tau$
Neutralinos	$\frac{1}{2}$	-1	$\tilde{B}^0,\, \tilde{W}^0,\, \tilde{H}_u^0,\, \tilde{H}_d^0$	$\tilde{\chi}_1^0,\, \tilde{\chi}_2^0,\, \tilde{\chi}_3^0,\, \tilde{\chi}_4^0$
Charginos	$\frac{1}{2}$	-1	$\tilde{W}^\pm,\, \tilde{H}_u^+,\, \tilde{H}_d^-$	$\tilde{\chi}_1^\pm,\, \tilde{\chi}_2^\pm$
Gluinos	$\frac{1}{2}$	-1	$\tilde{g}$	$\tilde{g}$
Gravitino	$\frac{3}{2}$	-1	$\tilde{G}$	$\tilde{G}$

Table 2.3.: Not yet discovered particles predicted by the MSSM, including the extended Higgs sector. The mixing in the first two generations is assumed to be negligible. The second column shows the so-called R-Parity, which will be explained in section 2.3.4. [33].

2.3.4. R-Parity

The superpotential for the MSSM reads as follows:

$$W_{\text{MSSM}} = \underbrace{\bar{u}\boldsymbol{y_u}QH_u - \bar{d}\boldsymbol{y_d}QH_d - \bar{e}\boldsymbol{y_e}LH_d}_{\text{Yukawa couplings for chiral supermultiplets}} + \underbrace{\mu H_u H_d}_{\text{Higgs mass term}} \quad , \tag{2.89}$$

with the chiral superfields as in Tab. 2.2 and $\boldsymbol{y_u}$, $\boldsymbol{y_d}$ and $\boldsymbol{y_e}$ being 3×3 matrices containing the Yukawa coupling parameters. In principle, there is no reason why there should not be

other terms like

$$W_{\Delta L=1} = \frac{1}{2}\lambda^{ijk} L_i L_j \bar{e}_k + \lambda^{'ijk} L_i Q_j \bar{d}_k + \mu^{'i} L_i H_u \tag{2.90}$$

$$W_{\Delta B=1} = \frac{1}{2}\lambda^{''ijk} \bar{u}_i \bar{d}_j \bar{d}_k, \tag{2.91}$$

with the generation indices $i = 1, 2, 3$. However, these terms would violate lepton number conservation (Eqn. 2.90) or baryon number conservation (Eqn. 2.91), respectively. As a consequence, if both $\lambda^{'}$ and $\lambda^{''}$ are not suppressed somehow, the proton's lifetime would be significantly shortened, which is not at all supported by experiment. To restore proton stability, a new quantum number, called R-parity, is introduced:

$$P_R = (-1)^{3(B-L)+2s}, \tag{2.92}$$

with B the baryon number, L the lepton number and s the particle's spin. With this definition, Standard Model particles are required to have R-parity of $+1$, while supersymmetric particles (sparticles) will yield an R-parity of -1. If R-parity is conserved, not only are terms like Eqn. 2.90 and Eqn. 2.91 forbidden, it also has other consequences:

1. Standard Model particles and sparticles do not mix with one another.
2. A supersymmetric particle has to decay into an odd number of other sparticles.
3. The lightest supersymmetric particle is stable.
4. In collider experiments, sparticles can only be created in pairs.

While all of these statements are phenomenologically of relevance, the third one is especially noticeable, as the lightest sparticle – if it is also electrically neutral – can serve as a candidate for dark matter, as was also mentioned before in the context of the lightest neutralino [33].

2.3.5. Supersymmetry Breaking

As has been mentioned in Sec. 2.3.2, Supersymmetry in its "original" form is constructed such that all fermions and bosons that belong to the same supermultiplet have the same masses and the vacuum energy is 0. The first would imply that those particles should have been discovered long ago, along with the Standard Model particles – which was definitely not the case, though. Consequently, Supersymmetry cannot be an exact symmetry, but rather broken by some means, i.e. the vacuum energy should be positive. However, the Supersymmetry breaking must not be too large in order not to destroy the cancellation of loop diagrams that solved the hierarchy problem. Therefore, the usual notion is that there is "soft" Supersymmetry breaking, introduced by some term added to the otherwise Supersymmetry-respecting Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}}. \tag{2.93}$$

There are several possibilities for the origin and form of $\mathcal{L}_{\text{soft}}$. One of them is to introduce some "hidden sector", which only interacts very weakly with the particles of the supersymmetric Standard Model via messenger particles. The Supersymmetry breaking would only enter the model via this hidden sector and otherwise Supersymmetry would be conserved. The symmetry breaking can also be spontaneous, similar to electroweak symmetry breaking, but as the generator of supersymmetric transformations Q is a fermionic

operator, instead of a massless Goldstone boson, there would be a massless fermion, a Goldstino. In analogy to the mass-generating mechanism for the electroweak boson, the Goldstino degrees of freedom could be absorbed by the gravitino, which would thereby obtain mass. The supersymmetric breaking models differ by their messenger mechanism. For example in case of supergravity models the Supersymmetry breaking is mediated by gravity (SUGRA). Typically, the gravitino is very heavy in such models, so that usually the lightest supersymmetric particle is the lightest neutralino. Another possibility is gauge mediated Supersymmetry breaking (GMSB), in which gauge interaction particles play the role of the messenger particles. In contrast to supergravity models, the gravitino is expected to be very light in this kind of model and is often assumed to be the lightest supersymmetric particle [2, 10, 33].

2.4. Direct Stau Production

With the lightest neutralino being a dark matter candidate in the considered scenario, the stau could also play a role in co-annihilation processes of neutralinos via the t-channel exchange of a stau. If the stau is relatively light, it could lead to a dark matter density that is compatible with cosmological evidence. [37, 38] In addition, if the neutralino and the stau are the lightest and next-to-lightest sparticles and the production of charginos and the heavier neutralinos is kinematically suppressed, the pair production of staus would be the dominant electroweak production mechanism of supersymmetric particles at the LHC.

2.4.1. Specifications of the Considered Model

The focus of the analyses presented in this thesis will lie on the search for the pair production process of two staus, as depicted in Fig. 2.4. The pair production occurs via a Drell-Yan process. The staus can be either mass eigenstate of $\tilde{\tau}_1$ and $\tilde{\tau}_2$, where it is assumed that the mixing angle is 0, i.e.

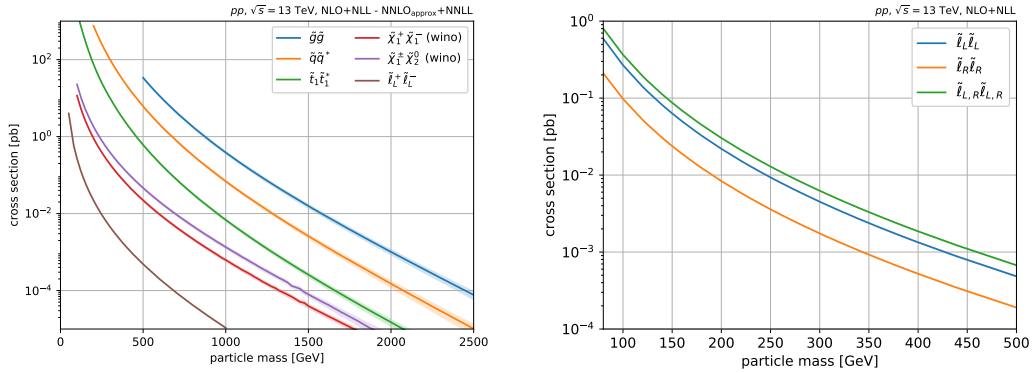
$$\tilde{\tau}_1 = \tilde{\tau}_L \tag{2.94}$$

$$\tilde{\tau}_2 = \tilde{\tau}_R. \tag{2.95}$$

Furthermore, both mass eigenstates are chosen to be mass-degenerate. In the course of this thesis stau masses between 80 and 560 GeV will be considered. The corresponding cross sections are in a range of 800 to 0.5 fb. The cross sections for slepton pair production in proton-proton collisions at a center-of-mass energy of 13 TeV are graphically shown in Fig. 2.3 and specifically listed for the considered stau pair production models in Sec. A. Compared to other sparticle pair production processes, slepton pair production has very small cross sections. Therefore, the presented analyses benefit greatly from the large integrated luminosity and center-of-mass energy in run 2.

The staus subsequently decay into one Standard Model τ-lepton and one lightest neutralino each. The lightest neutralino is taken to be the lightest supersymmetric particle and therefore, as R-parity should be conserved, it is stable. Moreover, the neutralino consists purely of the bino field without any admixture of winos and Higgsinos. As the neutralino interacts neither electromagnetically, nor via the strong force, it cannot be measured by the ATLAS detector (cf. Sec. 3.2). Therefore the neutralinos will cause a signature of missing transverse energy (cf. Sec. 4.3.5). The τ-leptons decay further, either into light leptons (electrons or muons) or into a number of hadrons, predominantly with either one or three charged mesons (referred to as 1-prong and 3-prong). Hence, one can identify three decay channels:

- LepLep channel: Both τ-leptons decay leptonically. This decay channel is very hard to pursue in an analysis, as the branching fraction for this decay is very low (in addition to the already very low production cross section).
- LepHad channel: One τ-lepton decays leptonically, the other into hadrons. The investigation of this channel also proved very challenging due to very large Standard Model backgrounds [45].
- HadHad channel: Both τ-leptons decay hadronically. This decay channel will be the main focus of this thesis.



(a) Sparticle production cross sections

(b) Slepton production cross sections

Figure 2.3.: Sparticle production cross sections in proton-proton collisions at 13 TeV in dependence on the sparticle mass. The left figure shows the predicted production cross section for several sparticles in comparison. The right plot illustrates the total cross section of slepton pair production in comparison with the separate cross sections for $\tilde{l}_L$ pair production and $\tilde{l}_R$ pair production. [39–44]

To summarize, the detector signature in question will be two hadronically decaying τ-leptons together with some amount of missing transverse energy.

The scenario described above is used as a simplified model. This means that any other supersymmetric particle, apart from the stau and the lightest neutralino, is assumed to be kinematically out of reach, so that the direct stau pair production process is the dominant and only supersymmetric process to occur. Furthermore, the decay of the staus into τ-leptons and LSPs is supposed to occur with a branching fraction of 100%.

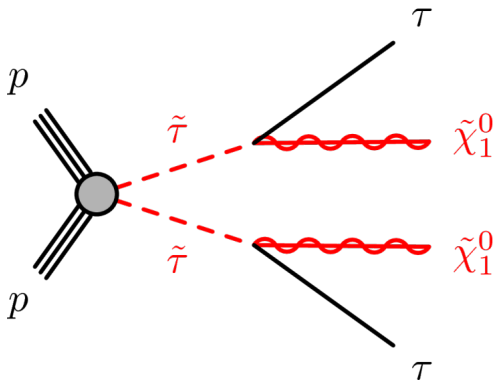


Figure 2.4.: After the direct production of two scalar τ-leptons each of them decay into a Standard Model τ-lepton and a lightest neutralino, which is assumed to be the LSP, and contributes to missing transverse energy. The particular decay channel depends on the decay modes of the two τ-leptons [46].

2.4.2. Previous Experimental Results

There have already been attempts to find direct stau pair production, e.g. at the Large Electron-Positron collider (LEP), which could exclude the pair production of $\tilde{\tau}_R < 90\,\text{GeV}$ with neutralino masses up to 80 GeV at 95% CL, as can be seen in Fig. 2.5 [47].

A first search for direct stau production with the ATLAS detector with the run 1 data ($20.3\,\text{fb}^{-1}$) of the LHC at 8 TeV center-of-mass energy could not exclude any stau and neutralino masses, but set upper limits on the cross sections [48].

In 2017, CMS published a paper with results with parts of the LHC run 2 dataset ($35.9\,\text{fb}^{-1}$) but also could not improve the exclusion limits on stau masses but only on the production cross section [49].

Two years later, in March 2019, CMS published the exclusion of stau masses for the combined eigenstates $\tilde{\tau}_L$ and $\tilde{\tau}_R$ up to 150 GeV for massless lightest neutralinos and set upper limits on the cross section for $\tilde{\tau}_L$ with a part of the LHC run 2 dataset ($77.2\,\text{fb}^{-1}$) [50].

In the same year, ATLAS published results based on the full run 2 dataset ($139\,\text{fb}^{-1}$), excluding stau masses between 120 GeV to 390 GeV at 95% confidence level for the combined $\tilde{\tau}$ eigenstates and also setting exclusion limits on the $\tilde{\tau}_L$ mass [51]. This analysis will be described in more detail in Chapter 6. In Chapter 7 a new search for stau pair production is presented which aims to improve the results from [51].

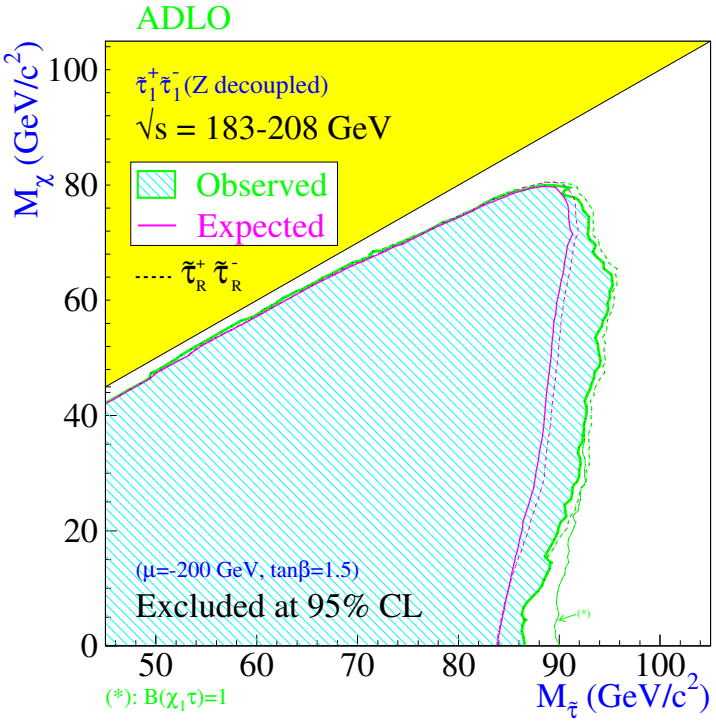


Figure 2.5.: Exclusion limits at 95% CL on stau and neutralino masses in the context of direct stau production at LEP. These limits are only derived for $\tilde{\tau}_R$. [47]

3. Experimental Setup

3.1. The Large Hadron Collider

As the masses of experimentally found particles pushed to higher and higher energy scales and indications pointed to new physics at even higher energy scales, it was necessary to construct a machine which made it possible to cover at least parts of these scales. This machine is the Large Hadron Collider (LHC) which stands in line with a long history of accelerators built at CERN, Geneva. The LHC's purpose is to accelerate and collide beams consisting of protons at a design energy of 7 TeV. With these collisions the existing understanding of particle physics can be probed and hints for new physics can be searched for and explored. However, before the protons can be injected to the LHC, they have to pass through a chain of pre-accelerators.

The protons are gathered from Hydrogen gas, which is stripped from its electrons and grouped into bunches of about 10^{11} protons before entering the first step of acceleration inside the LINAC2 machine. LINAC2 is a linear accelerator which uses radiofrequency (RF) cavities to accelerate protons so that their energy is increased to 50 MeV by the time they exit the machine. Quadrupole magnets are used to prevent the proton beam from dispersing. LINAC2 will be replaced by LINAC4 for run 3 of the LHC, which will probably start in 2021 [52].

After leaving LINAC2, the protons enter the Proton Synchrotron Booster. The Booster consists of four acceleration rings, in which the protons are accelerated to 1.4 GeV [53].

The next step in the acceleration chain is the so-called Proton Synchrotron (PS). The PS is a ring accelerator with a circumference of 628 m and at the time it was constructed (1959) it was the accelerator with the highest beam energy in the world. Today it not only serves as a part of the pre-accelerator chain of the LHC, but also supplies other machines and experiments with protons at an energy of 25 GeV [54].

Leaving the PS, the protons are injected to the Super Proton Synchrotron (SPS). With its almost 7 km circumference it is the first accelerator in the chain to be located underground. Originally it was built to provide collisions and host experiments itself, which e.g. lead to the discovery of the W- and Z-bosons in 1983. Today, it accelerates protons up to 450 GeV to supply the LHC and other experimental areas at CERN [55].

Finally, the protons enter the LHC to be accelerated up to their final energy of 6.5 TeV in run 2. The LHC consists of two separate beam pipes, in which the proton bunches are directed in opposite directions to be collided at four distinct points in the ring. The accelerator is located in a 26.7 km long tunnel in a depth between 45 and 170 m below the surface, which was previously housing the Large Electron Positron Collider (LEP). The entire acceleration chain is also depicted in Fig. 3.1. The acceleration inside the LHC is performed at one point of the ring (known as point 4), which hosts a complex of four modules containing four RF cavities each. These RF cavities operate at a frequency of

400 MHz, which corresponds to ten times the frequency of the proton bunches passing the accelerator module. The modules are placed inside a cryogenic system which enables operation in superconducting state. The remaining parts of the LHC are engineered so that the proton bunches are kept on a stable circular path and a sufficiently focused beam is sustained. The first is achieved by dipole magnets while the latter is accomplished with quadrupoles and higher multipole magnets.

In total there are 1232 dipole magnets built into the LHC, each of them with a length of 15 m. The magnets consisting of NbTi coils are superconducting and operating at 8 T. Cooling is ensured using superfluid helium at a temperature of 1.9 K. The dipole magnets around each of the beam pipes are placed inside the same "cold mass" and cryostat. The superconducting quadrupole magnets mentioned before are built-in as modules of about 5 m length and compensate the dispersion occurring inside the proton bunches by focussing in both x- and y-direction (in the transverse beam plane) [56].

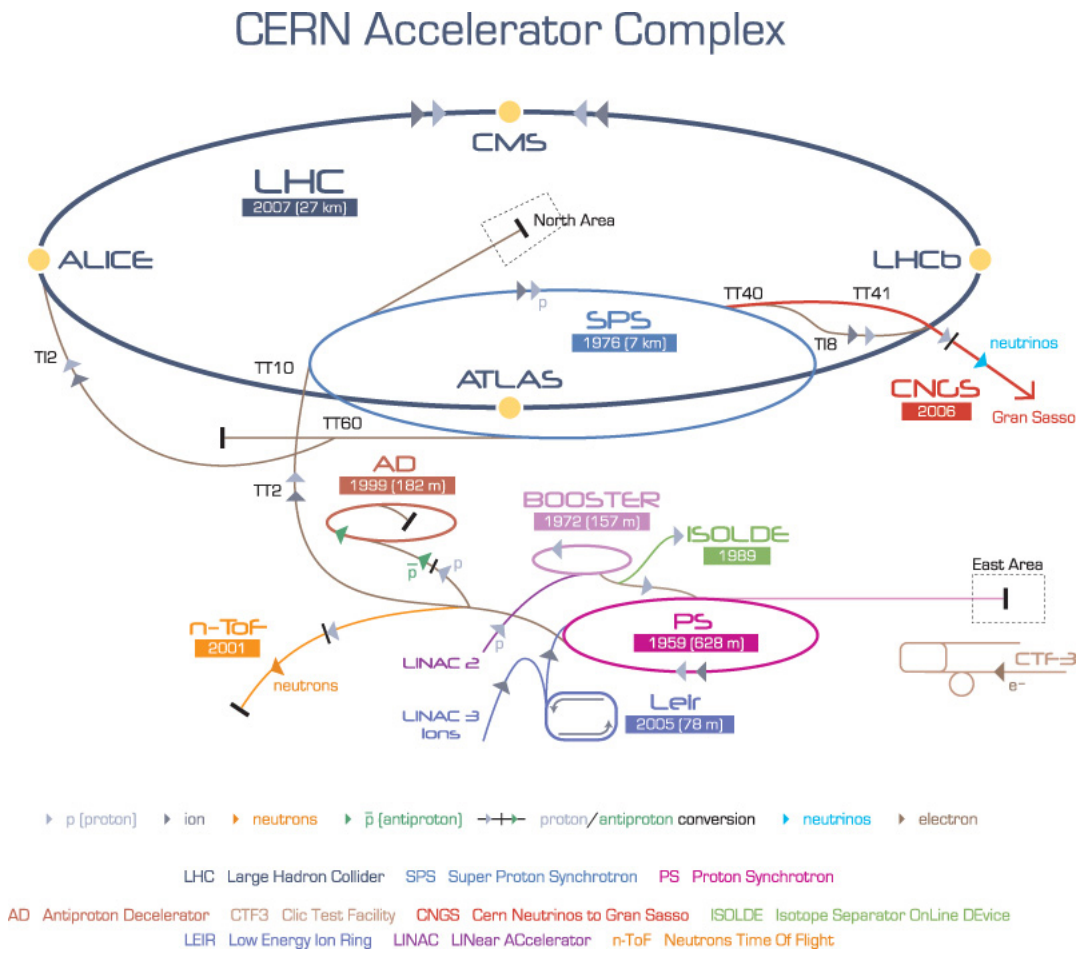


Figure 3.1.: A schematic image of the collider system located at the CERN site near Geneva. The largest accelerator ring is the LHC which is provided pre-accelerated protons by smaller accelerators like the SPS. There are four distinct locations, where the two proton beams can collide, each of them hosting a particle detector [57].

As was already mentioned, the two beams are brought to collision at four distinct points.

At each of these points a particle detector is located, which records the particles that are generated during the collisions. The proton bunches meet in intervals of 25 ns. The number of events generated in these collisions is given by

$$N_{\text{event}} = \sigma_{\text{event}} \cdot L, \tag{3.1}$$

with σ_{event} the cross section of the considered event type and L the machine luminosity. The machine luminosity is directly related to the properties of the proton beams:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F. \tag{3.2}$$

N_b is the number of protons per bunch, while n_b denotes the number of bunches per beam. Furthermore, f_{rev} is the rotation frequency, γ_r the relativistic Lorentz factor, ϵ_n the emittance of the beam in the transverse plane[1], β^* the beta function at the collision point and F a geometrical factor to account for the crossing angle of the beams at the interaction point [56].

The four experiments recording the proton-proton collisions are ATLAS [59], CMS [60], LHCb [61] and ALICE [62]. While ATLAS and CMS are multipurpose detectors, that are designed to perform a variety of measurements and searches, the LHCb detector is specialized at precision measurements of rare decays of B-hadrons and CP violation. As the LHC also is able to accelerate and collide lead ions, with ALICE there is a detector dedicated to the measurement of these kind of collisions. The ALICE experiment focuses on investigating the effects of QCD and exploring an exotic state of matter called quark-gluon plasma. The ATLAS detector will be discussed in more detail in the next section.

[1] The beam emittance characterizes the average (transverse) spread of a particle beam concerning position and momentum and is quantified through the standard deviation of a Gaussian beam profile [58].

3.2. The ATLAS Detector

The ATLAS detector is one of four detectors at the Large Hadron Collider. It is designed to measure a large variety of particles that emerge from proton-proton collisions, covering a wide range of possible measurements and searches. While the search for the Higgs boson was one of the most important benchmarks for the detector design, ATLAS is also capable of high-precision measurements of Standard Model particles, cross sections and couplings and of searches for e.g. new heavy gauge bosons, extended Higgs models, flavour changing neutral currents, extra dimensions and Supersymmetry. Its shape is that of a cylinder with a length of 44 m and a diameter of 25 m. The detector components are arranged in layers around the nominal interaction point, with dedicated detection systems not only present in the barrel but also in the endcaps of the cylinder. In the following, the ATLAS detector's structure and main properties and functionalities will be discussed. A schematic view of its interior structure is shown in Fig. 3.2.

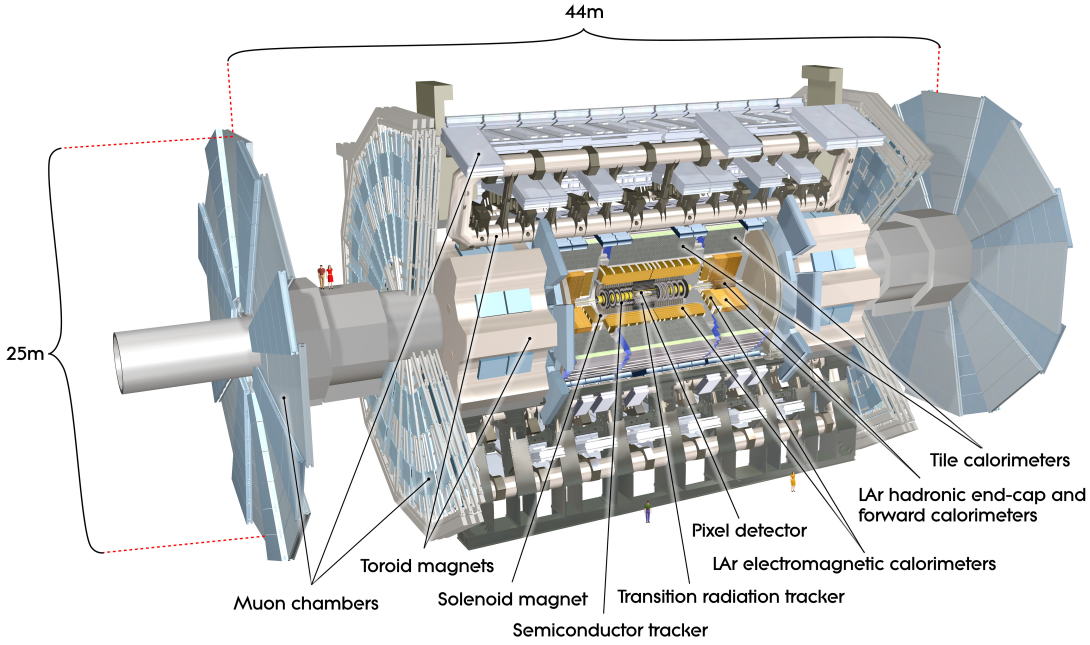


Figure 3.2.: The ATLAS detector is composed of several separate detector systems, which can be roughly divided into the inner detector (tracking of charged particles), the electromagnetic calorimeter (energy measurement of electromagnetically interacting particles), the hadronic calorimeter (energy measurement of strongly interacting particles) and the muon spectrometer (tracking detector for charged minimally ionizing particles) [63].

3.2.1. Coordinate System

The ATLAS detector's shape is cylindrical, with the proton beams entering through its bases. Per convention, the coordinate system is chosen such that the nominal interaction point is its origin and the beam axis corresponds to the z-axis. The xy-plane is perpendic-

ular to the beam direction, with the x-axis pointing towards the centre of the LHC ring and the y-axis upwards. The azimuthal angle ϕ is measured around the z-axis while the polar angle θ denotes the angle from the z-axis. A very commonly used property is the pseudorapidity, which is defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.3}$$

For massless particles, the pseudorapidity is identical to the rapidity y, which follows the definition

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right). \tag{3.4}$$

However, for relativistic particles (as they usually emerge from the hard proton-proton interactions) the mass is negligible and the pseudorapidity becomes a very good approximation. Another advantage of the pseudorapidity is that differences in η are Lorentz invariant under boosts in the longitudinal direction. Instead of the common polar coordinates ϕ and θ, the coordinate system of ϕ and η is therefore very often used with the ATLAS detector. The distance in $\eta\phi$-space is given as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.5}$$

Often, transverse quantities are used rather than the full 3- or 4-vectors. E.g. the transverse momentum p_{T} and the transverse energy E_{T} are defined in the xy-plane. The advantage of transverse quantities is given by the fact that the total momentum and energy in the transverse plane is well-known in contrast to the direction parallel to the beam axis. This allows for a better event reconstruction in the transverse plane in case undetectable particles are created in the event by exploiting energy and momentum conservation [59].

3.2.2. Tracking Systems (Inner Detector)

The innermost part of the ATLAS detector consists of three different systems, all of which serve the purpose of tracking the charged particles that are created in the collisions. The component closest to the beam tube consists of four layers of silicon pixel detectors. Subsequently, particles have to cross eight layers of silicon microstrip trackers (SCT), which are arranged so that all coordinates of the passing through particles can be measured. These components are not only present in the barrel region, but also in the end-cap regions. The pixels and the SCT cover the region of $|\eta| < 2.5$. They fulfil very high accuracy requirements and are crucial e.g. for vertex determination and identification of heavy quarks and τ-leptons. The third component is the transition radiation tracker (TRT), which consists of 4 mm straw tubes that are filled with a xenon-based gas mixture. In the barrel, straws with a length of 144 cm are oriented parallel to the beam axis, while in the end-caps the straws have a length of 37 cm and are arranged radially in wheels. The TRT covers a region of $|\eta| < 2.0$ and is especially useful for identifying electrons by transition radiation photons.

The combination of all three subsystems of the inner detector provides a high-accuracy tracking machinery, which can be used for very robust pattern recognition. The inner detector is surrounded by a 2 T solenoid magnet, which is used to bend the charged particles' trajectories. With the information of the bending radius and direction, the charges and momenta of the particles can be determined. A schematic picture of the inner detector structure in shown in Fig. 3.3 [59].

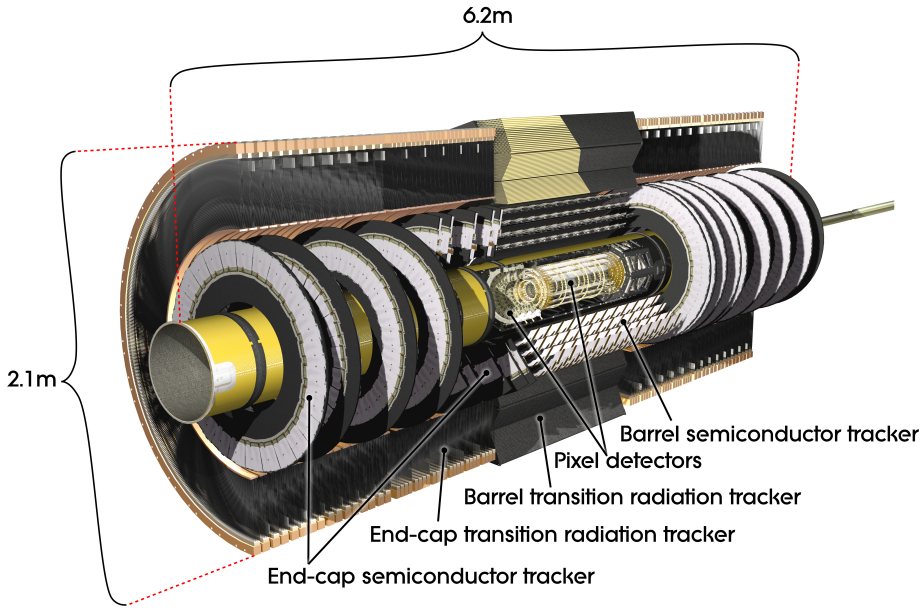


Figure 3.3.: Schematic view of the inner detector. The innermost part of the ATLAS detector contains three different tracking systems: The pixel detectors, the semiconductor tracker and the transition radiation tracker. [64]

3.2.3. Calorimeter Systems

The calorimeter system records the energy loss of particles that enter the detector components by causing electromagnetic or hadronic showers. Most particles that interact electromagnetically or via the strong interaction are expected to be stopped in this part of the ATLAS detector, with the exception of muons and possibly other minimally ionizing particles that are yet unknown. Therefore, the calorimeters have to be sufficiently thick, to contain the electromagnetic and hadronic showers and prevent punch-throughs to the muon system. This, together with a wide η-coverage of $|\eta| < 4.9$, enables a good measurement of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$, cf. Sec. 4.3.5). Roughly, the calorimeter system can be divided into two parts: The electromagnetic and the hadronic calorimeters, as can also be seen in Fig. 3.4 [59].

Electromagnetic LAr Calorimeter (ECAL)

The electromagnetic calorimeter possesses a very fine granularity and is able to measure electrons and photons with very high precision. It is split into a barrel part ($|\eta| < 1.475$) and two endcaps ($1.375 < |\eta| < 3.2$). The barrel is again divided into two half-barrels, leaving a small gap of 4 mm between them at $z = 0$. Similarly, the endcap parts are composed of two coaxial wheels - an outer wheel at $1.375 < |\eta| < 2.5$ and an inner wheel at $2.5 < |\eta| < 3.2$. These wheels are also called electromagnetic endcaps (EMEC). The active material used for the calorimeter is liquid argon (LAr). Accordion shaped absorber plates made of lead and kapton electrodes are used [59].

Hadronic Calorimeters (HCAL)

In the barrel region, the hadronic calorimeter consists of the so-called tile calorimeter, which does not only include the region directly outside the barrel part of electromagnetic LAr calorimeter ($|\eta| < 1.0$), but also two extended barrel parts ($0.8 < |\eta| < 1.7$). It is a sampling calorimeter, using polystyrene tiles as active material and steel as absorber material. The light created in the scintillating tiles is collected at the edges of each tile by wavelength-shifting fibres and subsequently guided to photomultipliers for read-out.

In the endcaps, the hadronic calorimeter is also a sampling calorimeter, but with liquid argon as the active material and with copper absorbers. The hadronic endcap calorimeter (HEC) consists of two wheels on either side of the barrel and has a coverage of $1.5 < |\eta| < 3.2$. The inner one of the two wheels is made of 25 mm copper plates, while the other consists of 50 mm copper plates.

In order to cover also regions of higher pseudorapidity and to reduce the radiation levels for the muon spectrometer, there is another calorimeter system, called forward calorimeter (FCAL) inserted inside the EMEC and HEC wheels. It consists of three modules: The first one is a copper calorimeter, optimized for measurements of electromagnetically interacting particles, while the second and the third module are tungsten calorimeters, dedicated to the measurement of hadronic interactions. Again, liquid argon is used as active material. The coverage of the FCAL is $3.1 < |\eta| < 4.9$ [59].

3.2.4. Muon Spectrometer

The muon system lying outside the calorimeter system provides tracking of those particles that can pass the calorimeters without depositing large amounts of energy. This is predominantly true for Standard Model muons, which are minimally ionizing in their interaction behaviour with the absorber materials. Similar to the inner detector tracking system, a magnetic field is applied in the muon spectrometer, so that the muons' trajectories experience bending force. The magnetic field is generated by three superconducting air-toroid magnets – one in the barrel region and two for the endcaps. The tracking system consists of three layers of muon chambers. Furthermore the muon spectrometer has its own trigger system. The structure of the ATLAS muon system is sketched in Fig. 3.5 [59].

Toroid Magnets

All three toroid magnets consist of eight coils that are arranged radially and symmetrically around the z-axis, with the endcap toroids being rotated by 22.5° with respect to the barrel toroid. The latter provides bending to muon trajectories in a region of $|\eta| < 1.4$, while the magnetic field of the endcaps governs the region $1.6 < |\eta| < 2.7$. In the region in-between, i.e. $1.4 < |\eta| < 1.6$ both fields overlap, which is therefore called transition region. The bending power of the magnets is described by the field integral $\int B d\ell$, with B the magnetic field component orthogonal to the muon direction and the integral computed along the muon trajectory between the innermost and outermost parts of the muon tracking system. The barrel toroid generates 1.5 to 5.5 Tm of bending power, while the endcaps' bending power is about 1 to 7.5 Tm. In the transition region, where the barrel and endcap fields overlap, the bending power is reduced [59].

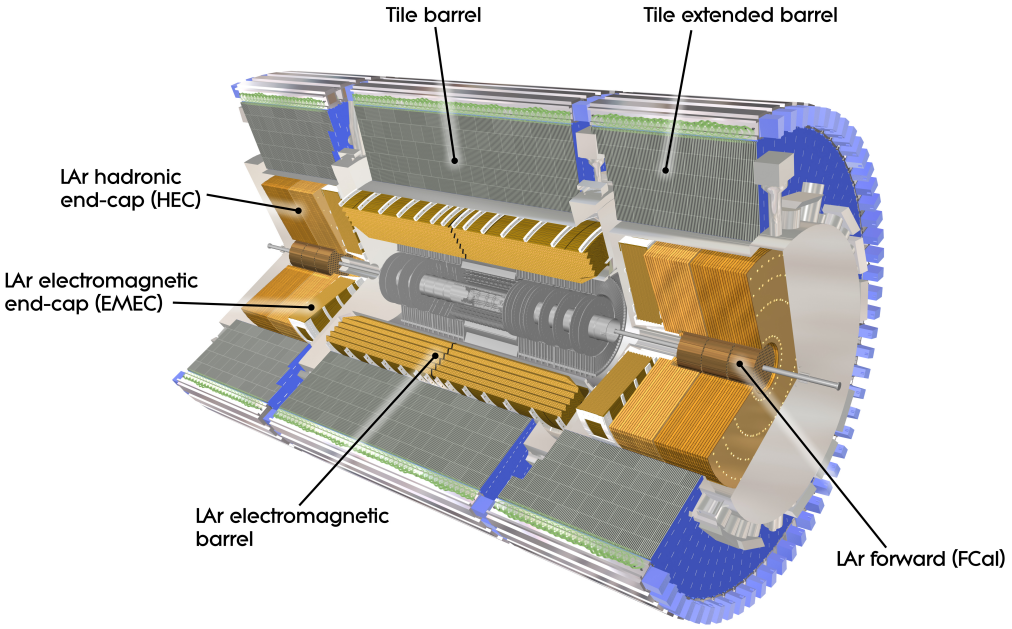


Figure 3.4.: Schematic view of the calorimeter systems. The inner layer of the calorimeter system is an electromagnetic LAr detector which also has its counterparts in the endcaps in the form of two coaxial wheels. The hadronic calorimeter is located outside the electromagnetic calorimeter - in the barrel with a tile calorimeter as well as in the endcap regions with LAr calorimeters. Another component, the forward calorimeter, is used for covering the region close to the beam tube [65]

Muon Chambers

There are two kinds of purposes the muon chambers have to fulfil: There are chambers designed for fast triggering and there are precision tracking chambers. In the barrel region, the trigger chambers are so-called resistive plate chambers (RPC), while in the endcaps thin gap chambers (TGC) take over this task. The trigger system covers an η-range of $|\eta| < 2.4$ and serves three purposes: The identification of bunch crossings, the determination of the muon coordinates in the direction orthogonal to the measurement of the precision tracking chambers and providing well-defined p_{T} thresholds. The precision tracking is performed by monitored drift tubes (MDT), which cover most of the η-range and cathode strip chambers (CSC), which are used at higher pseudorapidities [59].

3.2.5. Data Acquisition and Trigger System

Due to the bunch spacing in the LHC, the bunch crossings in the ATLAS detector occur every 25 ns for a time period of less than 1 ns. At the LHC's design instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ the average number of interactions in each bunch crossing is about 23. (During run 2 this number was varied for each data taking period. Cf. Sec. 4.1.) Together, this results in 10^9 interactions per second. This enormous number illustrates

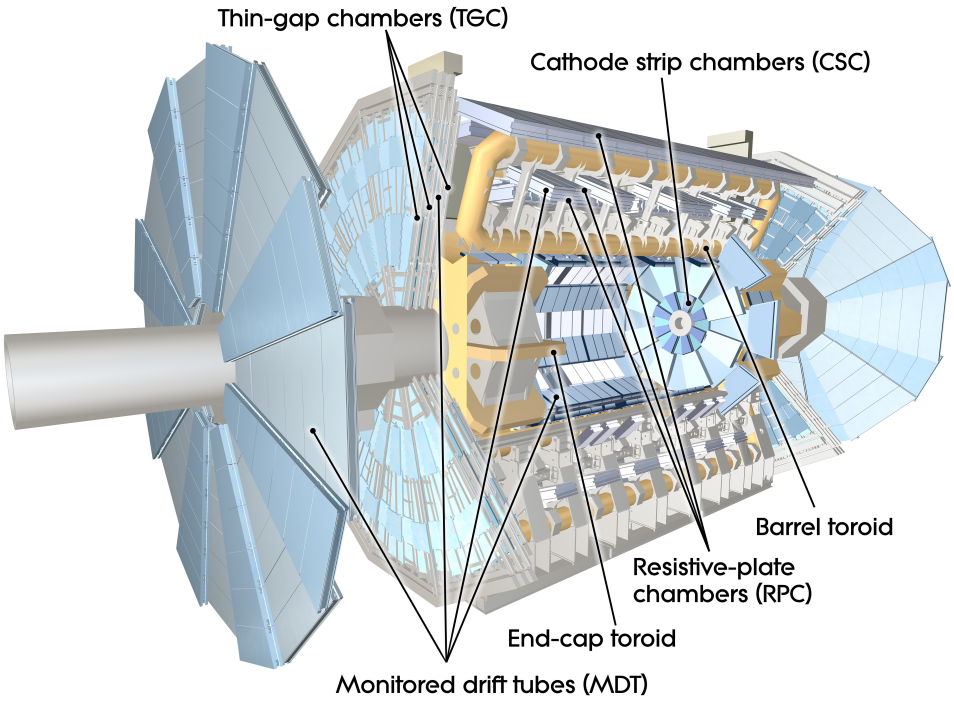


Figure 3.5.: ATLAS Muon Spectrometer: The muon system consists of multiple tracking detectors to reconstruct the trajectories of muons and other charged particles that might be able to pass through the calorimeter system. Similar to the inner detector, it is also placed inside a magnetic field, generated by a toroid magnet. [66]

why it is essential to run fast and selective trigger algorithms. In run 2 the trigger system consisted of two consecutive steps: The level 1 (L1) trigger and the high level trigger (HLT) [67, 68].

Level 1 Trigger (L1)

The level 1 trigger is hardware based and used to analyse data at a rate of 40 MHz and reduce it to 100 kHz for the HLT system within 2.5 μs. Relying on calorimeter signals and the information from the muon trigger chambers, the L1 trigger defines regions of interest (RoI) that serve as an input to the HLT. In addition, the L1Topo system allows to apply topological selections at L1 by considering kinematic information from multiple calorimeter and muon trigger objects, e.g. angular separation, invariant mass or the sum of the L1 jet momenta. The central trigger processor (CTP) provides the L1 trigger decision based on information from the L1 calorimeter trigger and the L1 muon trigger, as well as the L1Topo information. This system is schematically depicted inside the grey box in Fig. 3.6 [67].

High Level Trigger (HLT)

In run 1, the RoIs determined by L1 were first passed to a level 2 (L2) trigger before passing an event filter. For run 2, these two systems have been merged together into one single event processing HLT farm, which reduces complexity and code and furthermore allows for dynamic resource sharing of the algorithms. The HLT system performs an online reconstruction of physical objects within the RoIs passed to it by L1. Since run 1, the trigger reconstruction algorithms have been re-optimized, which reduced the differences of HLT and offline reconstruction algorithms. This proves especially true to the τ-lepton reconstruction, which lowered inefficiencies by more than a factor of 2. After the high level trigger system, the output rate is reduced to approximately 1 kHz compared to a rate of 100 kHz after L1 [67].

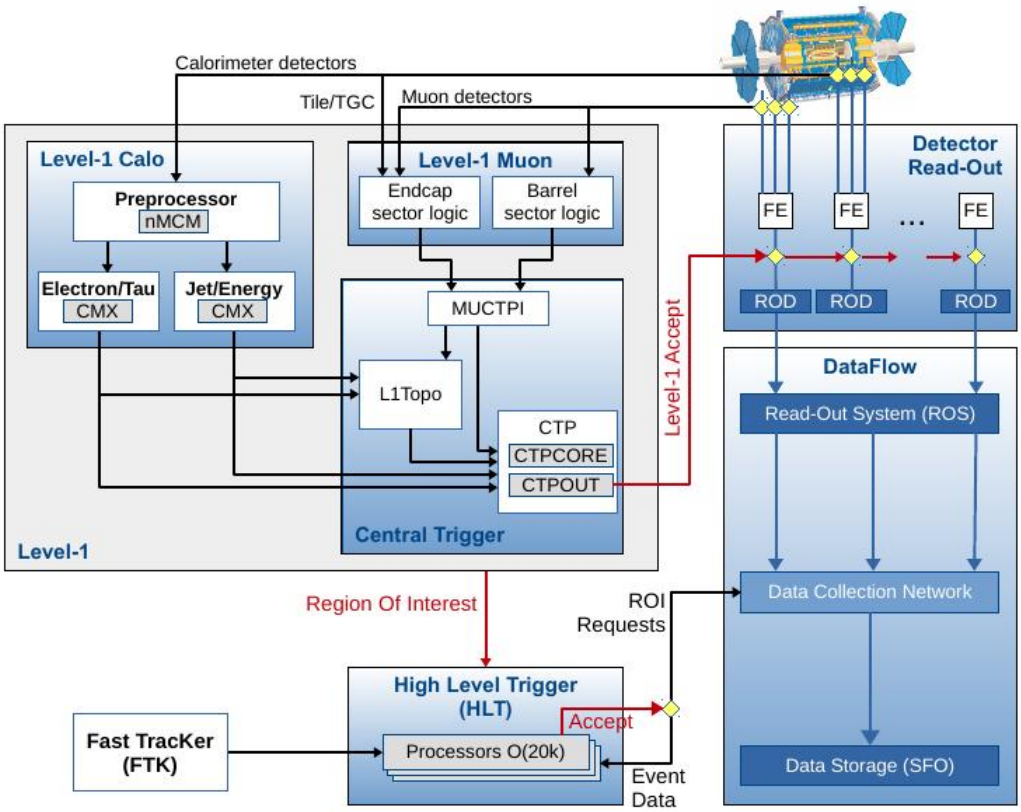


Figure 3.6.: ATLAS Trigger System: The level 1 trigger delivers regions of interest based on L1Topo, L1 calorimeter trigger and L1 muon trigger as input to the HLT system. The HLT performs an online reconstruction of physical objects based on these RoIs. The overall rate is reduced from 40 MHz to 1 kHz after HLT. [67]

4. Data and Monte Carlo Simulation

4.1. Data Taken at the ATLAS Detector in Run 2

Run 2 at the LHC lasted from 2015 to 2018 and was performed at a center-of-mass energy of 13 TeV. In total, the integrated luminosity recorded by ATLAS and passing all quality requirements sums up to $(139 \pm 2.4)\,\mathrm{fb}^{-1}$ [69]. The collection of data in terms of integrated luminosity over the whole data-taking period of run 2 can be viewed in Fig. 4.1. Out of this total integrated luminosity, $36\,\mathrm{fb}^{-1}$ were recorded in 2015-2016. In 2017, the integrated luminosity amounted to approximately $43\,\mathrm{fb}^{-1}$ and in 2018 to roughly $58\,\mathrm{fb}^{-1}$.

The number of interactions per bunch crossing was also different between the data taking periods. The average number of interactions per bunch crossing recorded by ATLAS is often referred to as pileup (μ). It corresponds to the mean of the Poisson distribution of the number of interactions per bunch crossing for each bunch and is calculated as

$$\mu = \frac{L_{\mathrm{bunch}} \cdot \sigma_{\mathrm{inel}}}{f_r}. \tag{4.1}$$

Here, L_{bunch} is the instantaneous luminosity per bunch crossing, σ_{inel} denotes the inelastic scattering cross section (which is taken to be 80 mb for collisions with 13 TeV center-of-mass energy) and f_r is the revolution frequency of the bunches in the LHC [69]. The pileup distribution for all years in run 2 as well as the total distribution can be seen in Fig. 4.2.

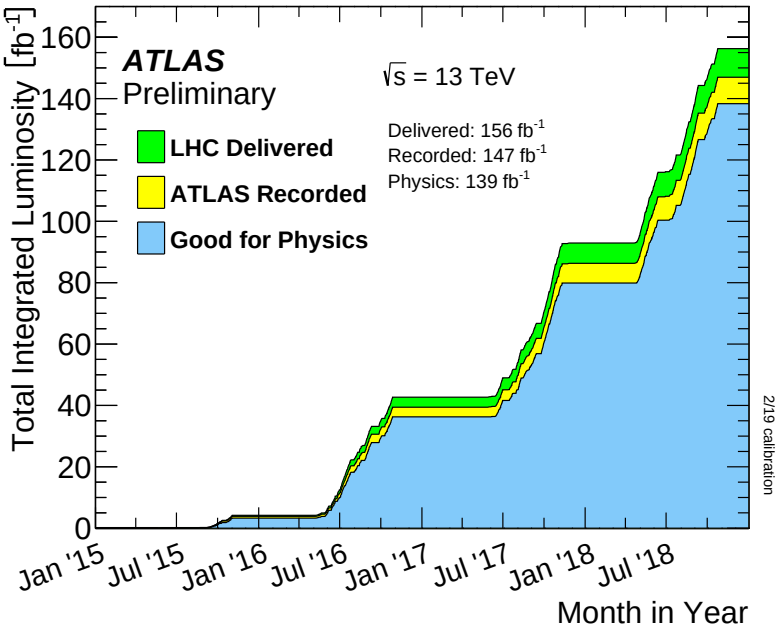


Figure 4.1.: Total integrated luminosity in run 2, delivered by the LHC (green), recorded by ATLAS (yellow) and complying quality criteria (blue, "good for physics"). [69]

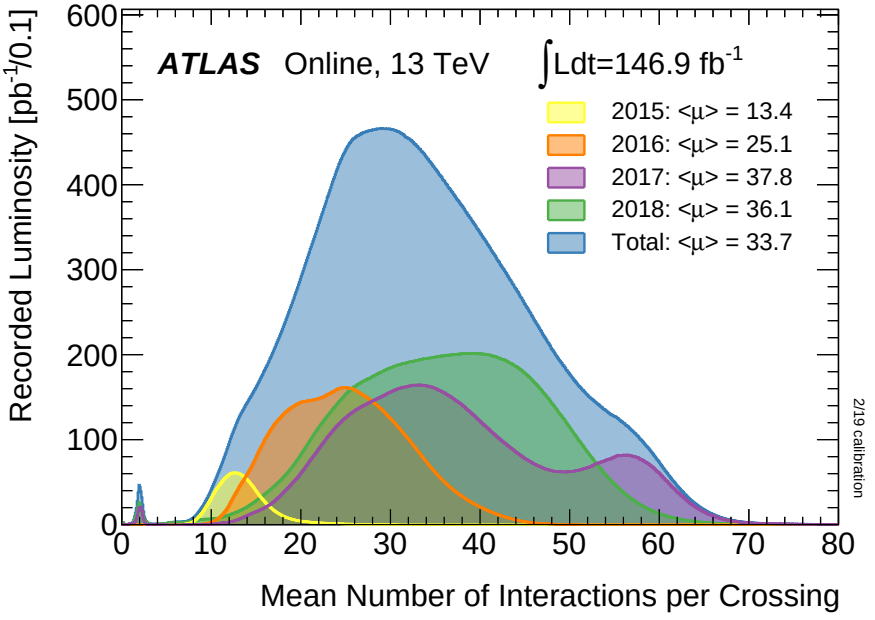


Figure 4.2.: Given here are the pileup profiles for all years of data taking in run 2, as well as a total pileup profile. [69]

4.2. Monte Carlo Simulations for Run 2

In order to avoid bias, the analysis of the collected data is first prepared based only on simulated data (or data in Standard Model dominated regions of phase-space like the estimation of multijet events, cf. Sec. 6.3.1 and Sec. 7.4.5). The simulated distributions are only compared to "real" ATLAS data once the analysis strategy is fully developed. Commonly, Monte Carlo simulations are used for this purpose.

4.2.1. Monte Carlo Method

The Monte Carlo method is a numerical tool to generate random events obeying a given probability density function (p.d.f.) in order to determine e.g. probabilities or other quantities in certain regions of phase-space. To shortly illustrate the Monte Carlo method, the case of a single random variable should be considered. At first, a sequence of random values $r_1, r_2, ...$ is generated that follow a uniform probability density function $g(r)$ between 0 and 1:

$$g(r) = \begin{cases} 1 & 0 < r < 1 \\ 0 & \text{otherwise} \end{cases} . \tag{4.2}$$

Afterwards, based on the random values r_i, a series of x_i are generated, which are distributed according to the desired p.d.f. $f(x)$. This step very often cannot be performed analytically, depending on the form of the p.d.f.. A possibility is von Neumann's acceptance-rejection technique. Consider a function $f(x)$ that can be surrounded by a rectangle given by $x_{\min}$, $x_{\max}$ and the maximum of the function $f_{\max}$: Then, the x_i are generated from the random values r_i such that $x_{\min} < x_i < x_{\max}$. Subsequently, a second independent set of random numbers u_i is generated, which follow a uniform distribution between 0 and $f_{\max}$. The key of von Neumann's technique is now to only accept x_i values for which $u_i < f(x)$. If not, the value x_i is rejected and the method is repeated. With this technique, the original p.d.f. $f(x)$ is approximated [70].

4.2.2. Event Generation

The Monte Carlo method, especially for more complex cases with more than one random variable, has many applications. In particle physics it is used to simulate e.g. scattering processes. The probability density function is then given by the according theory, e.g. in the form of a differential cross section. The Monte Carlo simulation algorithm then generates events with, for instance, different scattering angles or momentum vectors of particles. This is referred to as event generation. There is a large variety of event generation algorithms available in high energy physics, some of which will be mentioned in Sec. 6.3, when the particular simulated samples for the presented analysis are discussed [70].

For the purpose of simulating proton-proton collisions at the LHC, several components have to be considered, besides the "hard" scattering process, involving quarks and gluons inside the protons. The protons' internal structure is modelled with the help of parton distribution functions (PDF), which are external inputs to the simulation and rely on measurements performed on data of multiple experiments, such as detectors at the HERA accelerator, ATLAS, CMS and LHCb [71–73]. The PDFs document the probability to find

a parton at a certain fraction of its total momentum in a collision with a given momentum transfer [74].

Any colour charged particle can emit gluons with finite probability, which on their part may emit gluons or quarks, and so on. If particles emerging from the hard process are the source of these quark-gluon cascades, this is called final state radiation (FSR), in contrast to initial state radiation (ISR), which denotes colour charged particles emitted by one of the interacting partons. These QCD fragmentation processes occur on very small length scales and can therefore be calculated perturbatively and branching by branching. There are in fact two different approaches to deal with these gluon radiation cascades. One possibility is to calculate the matrix element of the hard process to higher and higher orders (where it is not possible to model soft or collinear radiation) or parton shower algorithms are used, that rely on iterative construction of the radiation branchings. Typically, with parton showers, in order to avoid soft and collinear divergences, a lower cutoff around 1 GeV is chosen, i.e. the simulation stops as soon as the scale of the QCD radiation branchings drop below this cutoff value. In practice, both approaches are often used together, e.g. by a procedure called "merging" [75–77].

As soon as all particles created in a collision, be it particles emerging directly from the hard process or due to ISR or FSR, move apart, they enter another QCD regime. At distances above ~ 1 fm asymptotic freedom does not govern the particles' behaviour anymore, instead they start to be confined in colour-neutral bound states. This stage is referred to as "hadronization" and occurs at the "hadronization scale", which is by construction at the lower cutoff of the parton shower. As this cannot be described perturbatively anymore, two phenomenological models are most commonly used in Monte Carlo generators today: The string model and the cluster model. The string model is a general term for models that describe hadronization processes as the breaking of "strings" that represent the colour connection between two quarks, for example: As the string is "stretched" the creation of a quark-antiquark pair can break it. The cluster model on the other hand relies on the principle of "preconfinement" i.e. the observation that partons in a shower form colour-neutral clusters with an invariant mass distribution that does not depend on the scale of the initial hard process. These clusters are interpreted as prototypes of hadrons. The hadrons created during hadronization might not be stable. Their decays then also have to be taken into account in the event simulation [75–77].

One further aspect to be considered by the simulation are the so-called underlying events, i.e. processes that involve the proton remnants that did not participate in the hard interaction [75–77].

The stages of event generation are summarized in Fig. 4.3.

4.2.3. Detector Simulation

Another important step is to simulate the detector signal for the outcome of the event generation, which is known as "core simulation". As there are also many random processes involved, such as ionization, this is done with the Monte Carlo method as well. A very widely used detector simulation algorithm is GEANT4 [78]. The outputs of these simulations are stored in terms of "hits", which are records of energy deposits with information on position and time. In a process called digitization, the hits are converted into digital detector responses or "digits". Typically, a digit is created when the voltage or current in

some read-out channel reaches a certain threshold. During digitization, also the fact that multiple protons can interact in each bunch-crossing is taken into account by overlaying the so far generated "signal" event with a collection of other events. The number of events to overlay is a function of the instantaneous luminosity that should be simulated, which is usually just an estimation of the actual pileup profile [74].

Both in the event generation and the detector simulation, additional information called "truth" is kept for each event. In the case of event generation, the truth records include the interaction processes and the involved particles together with their "identity" (or PdgID, cf. [76]). For the detector simulation, the truth information contains tracks and decays for certain particles. After digitization, the truth information along with the "ordinary" digitized detector simulation is processed by the trigger and reconstruction algorithms. Truth information allows, e.g. for performance checks of reconstruction algorithms, calibration and so-called fake background estimation and therefore is a necessary and beneficial addition to the simulations [74].

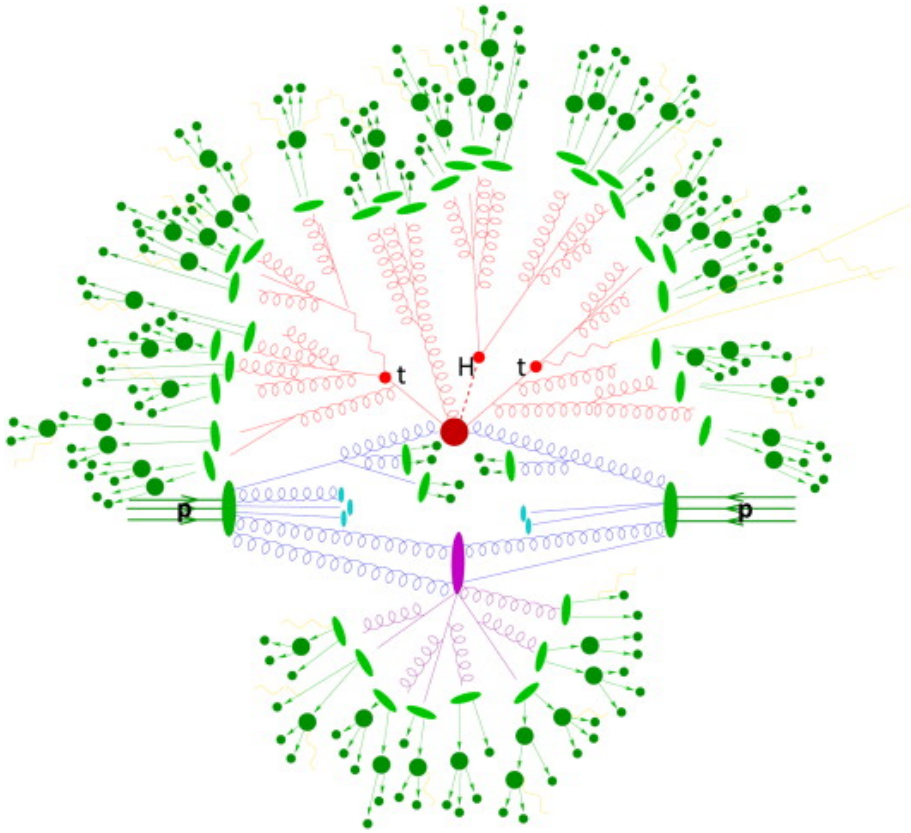


Figure 4.3.: Sketch of the various steps for generating a simulated proton-proton collision event. The incoming protons are depicted by the two green ellipses in the middle of the drawing. The red circle represents the hard process: Two gluons fuse to produce a four particle state ($t\bar{t}H$ + gluon) which is also drawn in red, as are the gluons and quarks that arise during the showering. The entire violet system represents an underlying event. The light green ellipses stand for the hadronization stage, while the dark green lines and circles depict hadronic decays. In this drawing, also electroweak radiation processes are considered, represented by the yellow lines. [79]

4.2.4. Event Weight

Monte Carlo simulated events can be equipped with weights, such that the distributions of variables are only predicted correctly if these weights are applied. There are different sources for event weights, such as the event generation process itself (depending on the generator), correction factors that have to be applied in retrospect. For example, the simulations of collision events inside the ATLAS detector are generated under a certain expectation for the number of interactions per bunch-crossing, i.e. pileup. This assumed pileup distribution is then corrected to correspond to the measured pileup in data. The correction is applied to the simulated events is form of a weight, w_{pileup}. Another type of correction that is taken into account by weighting the simulated events are so-called scale factors which should account for differences in performance of object reconstruction and identification algorithms (cf. Sec. 4.3) in simulation as compared to data. (Similar to the offline reconstruction, there are also scale factors to account for differences in the performance of triggers in simulation and data.) Here, these corrections are summarized as the weight $w_{\text{reconstruction}}$. For unweighted events, e.g. for data, the number of events for some given process is obtained like

$$N_{\text{process}} = \sigma_{\text{process}} \cdot \mathcal{L}, \tag{4.3}$$

with σ_{process} being the total cross section for the considered process and $\mathcal{L}$ the integrated luminosity. For weighted events, however, each of the individual weights for the i^{th} selected event has to be taken into account in addition to the cross section and the luminosity:

$$w_i = \frac{\sigma_{\text{process}} \cdot \mathcal{L}}{\sum\limits_{j \in \text{generated}} w^j_{\text{generator}}} \cdot w^i_{\text{generator}} \cdot w^i_{\text{reconstruction}} \cdot w^i_{\text{pileup}}. \tag{4.4}$$

The number of events for the considered process is then simply given as

$$N^{\text{weighted}}_{\text{process}} = \sum_i w_i. \tag{4.5}$$

The uncertainty on this number is given by the square root of the sum of squared weights:

$$\Delta N^{\text{weighted}}_{\text{process}} = \sqrt{\sum_i w_i^2}. \tag{4.6}$$

4.3. Object Reconstruction and Identification

Following the digitization step, the events are subjected to the HLT and afterwards to the reconstruction algorithms. A brief overview of the HLT was already given in Sec. 3.2.5. The next sections will outline the reconstruction and identification algorithms for different physics objects that are used in the analysis presented in this thesis.

4.3.1. Electron Reconstruction and Identification

The reconstruction of electrons (and photons) relies predominantly on identifying clusters of energy deposits in the calorimeter systems and tracks in the inner detector.

Topo-Clusters

For the first part, so-called topological clusters or topo-clusters are reconstructed in an iterative procedure. The seeds for the topo-clusters are cells inside the calorimeters (both the electromagnetic and hadronic calorimeters), which pass a certain significance criterion: $|\zeta_{\text{cell}}^{\text{EM}}| \geq 4$, with

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise, cell}}^{\text{EM}}}. \tag{4.7}$$

This requirement ensures that the energy deposit $E_{\text{cell}}^{\text{EM}}$ in the seed cell is sufficiently larger than the expected noise from electronics and pileup $\sigma_{\text{noise,cell}}^{\text{EM}}$. The procedure continues by selecting all neighbouring cells to the seed cells whose significance lies above a threshold $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$, which then provide the seeds for the next iteration. If two of these proto-clusters happen to share a cell with sufficient significance above noise, the clusters are merged. On the other hand, if the resulting cluster exhibits two or more local maxima of energy deposits, i.e. $E_{\text{cell}}^{\text{EM}} > 500\,\text{MeV}$ together with at least four neighbouring cells with lower energy deposits, the proto-cluster is split. In the final stage of building topo-clusters, only clusters with $E_{\text{cluster}}^{\text{EM}} > 400\,\text{MeV}$ in the electromagnetic calorimeter cells are selected. Although the building of topo-clusters starts with both the signals of the electromagnetic and the hadronic calorimeter, for the reconstruction of electrons and photons only the energy measured in the cells of the electromagnetic calorimeter are used. The fraction of a total topo-cluster's energy that was deposited in the electromagnetic calorimeter is expressed by the quantity f_{EM}. In order to suppress pileup induced clusters, a preselection cut of $f_{\text{EM}} > 0.5$ is placed [80].

Track Reconstruction

Besides the topo-clusters in the calorimeter, another key element to the reconstruction of electrons are tracks in the inner detector. A reconstruction algorithm first assembles clusters based on the hits in the pixel and SCT detector components. These clusters are then converted into three-dimensional measurements called space-points. Following this, the space-points serve as the input for a pattern recognition algorithm and a subsequent fit under the hypothesis of the tracks originating from a charged pion. If the fit under the pion hypothesis fails, an electron hypothesis is used instead. In order to better take account for the energy loss of charged particles inside material, an additional fit strategy, called

Gaussian-sum filter (GSF) method [81] is used. In this method, the material-induced energy loss is modelled by six Gaussian functions, while up to twelve Gaussian functions describe the track parameters. Of special interest considering the curvature of the track (induced by the magnetic field of the inner solenoid) are the transverse impact parameter d_0 and its uncertainty $\sigma(d_0)$, and their ratio, which is referred to as transverse impact parameter significance. The direction of the curvature furthermore determines the sign of the electric charge of the particle that caused the track. Ultimately, the inner detector tracks are matched to the topo-clusters in the electromagnetic calorimeter [82].

Superclusters

In a third step, so-called superclusters are reconstructed from the topo-clusters in the calorimeter. The topo-clusters are first sorted according to their transverse energy E_{T}. Each of the topo-clusters is tested for its use as a seed for the supercluster building algorithm. In order to be chosen as supercluster seed, a topo-cluster has to have a transverse energy of at least 1 GeV and be matched to a track in the silicon detectors with at least four hits. Following the seed identification, so-called satellite clusters are determined. A topo-cluster is considered to be a satellite cluster to a seed if it lies inside a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycenter. This is mostly the case for secondary showers induced by the same electron. However, a topo-cluster is also classified as satellite if it is within $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ around the seed's barycenter and the track it is matched to is identical to the best-matched track for the seed cluster. Together with their satellites, the seeds then form superclusters. After the superclusters are determined, they are once again matched to the tracks in the inner detector [80].

Energy Calibration

The superclusters serve as electron candidates for the identification procedure, after a certain calibration scheme to the electron energy, estimated from the energy deposits in the electromagnetic calorimeter, is applied. This calibration aims to correct for differences between data and simulation as well as to adjust the absolute and relative energy scales between the calorimeter layers. The estimated energy of the electrons is derived via a multivariate regression algorithm, trained on simulated events, which in addition optimizes the energy resolution. Local non-uniformities in the calorimeter response, e.g. from geometric effects at module boundaries, are also taken into account [80, 83].

Electron Identification and Isolation

Finally, the reconstructed and calibrated electron candidates are subjected to an identification procedure, in order to improve the purity of the selected prompt electrons. The main backgrounds "imitating" an electron signature originate from hadronic jets (cf. Sec. 4.3.3), converted photons and non-prompt electrons that are created in heavy-flavour quark decays. To achieve this, a likelihood discriminant is constructed based on variables related to the properties of the primary electron track in the inner detector, the lateral and longitudinal development of the electromagnetic shower in the calorimeter, and to the matching of tracks and clusters. The likelihoods L_S and L_B for the candidate to originate from signal (i.e. be a prompt electron) and from background, respectively, are calculated via

probability density functions P, which are obtained by the smoothing of histograms of n of the input variables:

$$L_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{S(B),i}(x_i). \tag{4.8}$$

The discriminant is constructed as the natural logarithm of the ratio of L_S and L_B. Based on this discriminant, three working points with different selection efficiencies are optimized: Loose ($\approx 93\%$), medium ($\approx 88\%$) and tight ($\approx 80\%$). The lower efficiency for the medium and tight working points corresponds in return to a better rejection of background.

Another property of reconstructed electrons of importance for analyses is the isolation, as this poses another way of discrimination to hadronic objects and non-prompt electrons. Usually, isolation criteria allow only for a certain amount of deposited energy in clusters or momenta in tracks that lie within a cone of typically $\Delta R = 0.2$ around the reconstructed electron cluster or the primary electron track, respectively. There are two kinds of working points that can be defined with these requirements: Either the efficiency is kept at a certain value or the cuts on the isolation variables is fixed. The so-called Gradient working point is set to be at an efficiency of 90% at a transverse momentum of 25 GeV and 99% at 60 GeV. The isolation working points referred to as HighPtCaloOnly, Loose and Tight rely on fixed cut values on the calorimeter or track isolation variables and thus have variable efficiencies [80].

4.3.2. Muon Reconstruction and Identification

The components of the muon spectrometer are the key elements in reconstructing muons. The muon spectrometer based reconstruction can then be combined with information from other sub-detector systems, such as the inner tracking detector.

Reconstruction in the Muon Spectrometer

The muon spectrometer reconstruction is an interplay of the signals of the different sub-systems. In general, so-called segments are constructed based on a hit pattern search inside each muon chamber. Subsequently, a fit of segments in different layers provides the muon track candidates. The fitting algorithm starts with segments in the middle layer of the detector as seeds because of the higher hit multiplicity available there, and later-on includes segments from the outer detector layers as well. Selection criteria based on the hit multiplicity and fit quality are applied on the segments before being matched to form track candidates in the fitting procedure. At least two matching segments are needed to build a track candidate with the exception of the magnetic transition region, where one high quality segment is sufficient. Because the same segments can in principle be used to construct multiple track candidates, an overlap removal is performed afterwards [84].

Combined Reconstruction

Other sub-detector systems may be taken into account for muon reconstruction as well. Depending on the used detector components, four different reconstructed muon object types can be defined:

- Combined muons (CB): First, the reconstruction of tracks is performed independently in the muon spectrometer and in the inner tracking system of ATLAS. A combined track is then obtained by performing a global refit including hits from both detector components. Usually, an outside-in pattern recognition technique is used, in which the reconstructed muon track candidates in the muon spectrometer are extrapolated inwards and matched to a track in the inner detector.
- Segment-tagged muons (ST): In this case, a track found in the inner detector is extrapolated to the muon spectrometer system and is defined as a segment-tagged muon object if it can be associated with at least one segment in the MDT or CSC subsystems. This type of muon object is useful in the event of the muon traversing only one layer of muon chambers, which can be due to low detector acceptance in that region or a low transverse momentum of the particle.
- Calorimeter-tagged muons (CT): In regions of reduced acceptance of the muon spectrometer due to cabling and service for the inner parts of the ATLAS detector, muons identified only through an inner detector track matched to an energy deposit in the calorimeter are used to compensate. For this purpose, the energy deposit in the calorimeter has to be compatible with that of a minimally ionizing particle.
- Extrapolated muons (ME): In order to make up for the region not covered by the inner detector $2.5 < |\eta| < 2.7$, extrapolated muon candidates are identified based only on the muon spectrometer track, which has to be compatible with originating from the interaction point.

If two kinds of muon candidates share the same track in the inner detector, CB muons are preferred over ST muons, which in turn are favoured over CT muons. Overlap with ME muons is resolved by placing requirements on the fit quality and number of hits [84].

Muon Identification

Objects wrongly classified as muons mainly originate from pion and kaon decays. These decays usually show a "kink" in the reconstructed tracks in the inner detector leading to a poor fit quality. This can be exploited to reduce such backgrounds by placing requirements on the normalized χ^2 of the combined track fit. Furthermore, decaying charged mesons lead to a difference in the momentum measurement in the inner detector and the muon spectrometer. Two variables that characterize this difference are the so-called q/p significance, which is defined as the absolute value of the charge to momentum ratio difference in the inner detector and the muon spectrometer, divided by the combined uncertainty, and ρ', which is the absolute value of the difference between the transverse momenta measured in the inner detector and the muon spectrometer, divided by the transverse momentum of the combined track. To ensure a good suppression of pion and kaon background as well as a reliable momentum measurement, quality requirements based on these variables are placed. In total, four different identification working points are defined:

- Loose: Designed to maximize the selection efficiency. It incorporates all types of reconstructed muons. The efficiency ranges from 96.7% for muons with low p_{T} to 98.1% for high p_{T} muons.
- Medium: The medium working point minimizes systematic uncertainties from reconstruction and calibration and only utilizes CB muons with a least 3 hits in no

less than 2 MDT layers and ME muons with hits in at least 2 MDT or CSC layers in a region of $2.5 < |\eta| < 2.7$. The q/p significance is set to < 7 together with further compatibility criteria on the momentum measurements in the inner detector and the muon spectrometer. The efficiency of this working point is computed to be 95.5% for low p_T ranges and 96.1% for higher p_{T} values.

- Tight: The tight working point maximizes purity and considers only CB muons with hits in at least two layers in the muon spectrometers. The normalized χ^2 is required to be smaller than 8. Still, muons are selected at an efficiency of 89.8% at low p_{T} and 91.8% at a higher p_{T} range.
- High-p_{T}: This is a dedicated working point which maximizes the momentum resolution for tracks with a very high transverse momentum above 100 GeV [84].

Muon Isolation

Muons which are created directly in the "hard" interaction process very often appear separately from other particles in the detector, whereas muons from semileptonic hadron decays mostly are embedded in jets (cf. Sec. 4.3.3). To discriminate these two muon sources, the activity in the detector close to the reconstructed muon is quantified by two isolation variables. The so-called track-based isolation variable $p_{\mathrm{T}}^{\mathrm{varcone30}}$ is defined as the scalar sum of all transverse momenta of tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ inside a cone around muon. The size of this cone is determined in dependence of the transverse momentum of the muon p_{T}^{μ} according to the following formula:

$$\Delta R = \min\left(10\,\mathrm{GeV}/p_{\mathrm{T}}^{\mu}, 0.3\right). \tag{4.9}$$

The calorimeter-based isolation variable $E_{\mathrm{T}}^{\mathrm{topocone20}}$ is determined as the sum of the transverse energy of all topo-clusters (cf. Sec. 4.3.1) within a cone of size $\Delta R = 0.2$ around the muon, taking corrections due to pileup effects into account. Several isolation working points are defined based on these two isolation variables. E.g., the FixedCutLoose working point requires for the track-based isolation $p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.15$ and for the calorimeter based isolation $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu} < 0.3$ and has close to 100% selection efficiency at transverse momenta $\gtrsim 30\,\mathrm{GeV}$ [84].

4.3.3. Jet Reconstruction and Identification

In ATLAS, the anti-k_t algorithm is used for jet reconstruction, which is an iterative combination procedure of jet constituents. It is related to the Cambridge-Aachen and the k_t algorithm, but has the advantage of resulting in jets with circular boundaries while being collinear and infrared safe. The iteration condition relies on a parameterization of the distance between the jet constituents, which includes a radius parameter R [85].
For the ATLAS jet reconstruction this radius parameter is usually chosen to be 0.4. In general, multiple types of four-vector objects can be used as input to the anti-k_t algorithm, resulting in different kinds of reconstructed jets. Topo-clusters built from calorimeter cells, which are corrected to the EM scale[1], were the standard jet seeds for ATLAS analysis until

[1] The detector response for electromagnetically interacting particles (electromagnetic scale) is different in ATLAS than that for hadrons (hadronic scale). As a consequence, calibrations are applied to the calorimeter cells to even out these differences. This can be either achieved by a global method in the context of jets or in a local method for separate topo-clusters (cf Sec. 4.3.4).

shortly after the end of run 2. In addition, these topo-clusters are required to originate from the primary vertex. The primary vertex is identified as the origin of tracks with the largest sum of squared track momenta. As the position of the the primary vertex differs for each event, a so-called origin correction is applied to the topo-clusters. The resulting jets of these reconstruction inputs are referred to as EMtopo jets.

After run 2, the jet reconstruction was further improved by introducing a way of incorporating more detector information from different subsystems. This so-called particle flow (PFlow) algorithm makes use of both measurements of the inner tracking detector and the calorimeter to provide input to the anti-k_t algorithm. The energy deposited by charged particles in the calorimeter is thereby replaced inside the topo-clusters by the matching momentum measurement delivered by the inner tracking systems. With PFlow jets, the reconstruction efficiency was improved as well as the energy and angular resolution compared to EMtopo jets [86].

In the course of this thesis, both types of jets will be used.

Jet Energy Scale and Jet Energy Resolution Corrections

Several effects lead to a jet signal different to what is expected by simulations. The jet energy scale (JES) calibration aims to compensate for that and restore the measured jet energy to that of simulated jets at truth level. The first calibration step is a correction for pileup effects. Additional proton-proton interactions during the same bunch crossing (in-time pileup) or a nearby bunch crossing (out-of-time pileup) cause extra energy deposits in the detector. Therefore, a correction using the jet area and the transverse momentum density of an event is performed as well as a correction for residual effects which is estimated via Monte Carlo simulations in dependence on the number of interactions per bunch crossing and the amount of reconstructed primary vertices per event. Another correction obtained by the comparison with Monte Carlo simulation is an absolute calibration of the energy and direction of the jets. Also, a dedicated calibration is performed to improve the transverse momentum resolution for jets and reduce fluctuations in the detector response that originate from the energy distribution inside the jets as well as the nature of its constituents. As a last step, a in situ calibration is conducted to correct for remaining differences in between data and simulation using well-measured reference particles such as photons and Z-bosons [86].

The resolution of the jet energy (JER) is an important input for many analyses involving jets, but also for the determination of the missing transverse momentum (cf Sec. 4.3.5). One can divide the origin of the jet energy resolution roughly into three different categories: Noise, stochastic origins and a constant term. The noise term is caused by electronic noise in the detector as well as pileup and is therefore independent of the jet to be measured. The stochastic component takes into account statistical fluctuations in the amount of deposited energy while the constant term considers fluctuations that pose a constant fraction of the jet transverse momentum, like energy deposits in non-active material and non-uniform response in the calorimeter. Besides having a good knowledge of the jet energy resolution, it is also necessary to make sure that the resolution in data and simulation are in agreement. To achieve this, a smearing procedure is applied [86].

Identification of Bottom-Quark-Induced Jets

Discriminating jets originating from b-hadrons against other kinds of jets is a useful tool for many ATLAS analyses. It relies mainly on the reconstructed tracks of charged particles in the inner detector and the jets associated to them. Several discriminants are identified based on track impact parameters and compatibility likelihoods of tracks with the reconstructed primary vertex. In addition, as muons can occur in b-hadron decays. a soft muon tagger is employed. Another approach is to identify secondary vertices with the help of vertex based taggers. All these discriminants serve as inputs to multivariate techniques, called high-level b-tagging. In this thesis, two high-level b-taggers have been used. One relies on a boosted decision tree discriminant, called MV2, while the other makes use of a deep neural network, referred to as DL1 [87].

4.3.4. Tau Reconstruction and Identification

The reconstruction of hadronically decaying τ-leptons is based on the collection of reconstructed jets in each event. Most hadronic τ-decays involve either one (1-prong) or three (3-prong) charged mesons. Leptonically decaying τ-leptons are, however, reconstructed as muons or electrons, respectively.

Tau Reconstruction

Jets, that are considered as τ-candidates, are required to have a transverse momentum above 10 GeV and must be in a region of $|\eta| < 2.5$, with the exclusion of the region between the barrel and the endcaps $1.37 < |\eta| < 1.52$. τ-vertices are identified as those track vertices with the largest fraction of momentum of tracks, which are located within $\Delta R < 0.2$ of the jet. Besides using the reconstructed jet and the determining τ-vertex, quality requirements are set to the τ-candidate's tracks. The transverse momentum of each track should be at least 1 GeV and must include a minimum number of hits in the tracking detector. The minimum distance of the track to the τ-vertex in the transverse detector plane d_0 must be smaller than 1 mm, while the shortest distance in the longitudinal plane has to fulfil $|\Delta z_0 \sin\theta| < 1.5\,\text{mm}$, with θ being the polar angle and z_0 the point of closest approach along the z-axis. The direction in terms of η and ϕ of the τ-candidate is obtained from the τ-vertex as origin and the topo-clusters inside the jet seed. The energy of the τ-candidate has to be derived through a dedicated calibration scheme [88].

Tau Energy Scale Calibration

Similar to the energy correction for jets, the τ-energy is calibrated to match the average energy carried by the measured τ-decay products in simulation. The τ-energy scale (TES) consists of several parts, the first being the so-called baseline calibration. The correction is applied to the sum of the energy of all topo-clusters in a range of $\Delta R < 0.2$ around the τ-candidate. This energy (E_{LC}) already has been subject to a local hadronic calibration (LC, cf Sec. 4.3.3) and is further corrected for energy deposited by pileup (E_{pileup}):

$$E_{\text{calib}} = \frac{E_{\text{LC}} - E_{\text{pileup}}}{\mathcal{R}(E_{\text{LC}} - E_{\text{pileup}}, |\eta|, n_p)}, \tag{4.10}$$

with $\mathcal{R}$ being the detector-response calibration factor, which is the Gaussian mean of the distribution $(E_{\mathrm{LC}} - E_{\mathrm{pileup}})/E^{\mathrm{vis}}_{\mathrm{true}}$, derived for each the 1-prong ($n_p = 1$) and the 3-prong ($n_p = 3$) case and in dedicated $|\eta|$-bins. $E^{\mathrm{vis}}_{\mathrm{true}}$ is the energy of the τ-decay products in simulation excluding neutrinos. Although the baseline calibration shows a good τ-energy resolution (TER) for higher values of transverse momenta, its performance degrades for lower p_{T}. Therefore an additional method is employed to improve the resolution also in the lower parts of the p_{T}-spectrum. With the "Tau Particle Flow" (TPF) algorithm, all individual charged and neutral hadrons in τ-decays are reconstructed, thus significantly improving the energy resolution. In order to benefit from both methods, the baseline calibration and the TPF method are used together with an additional multivariate technique called boosted regression tree (BRT) exploiting calorimeter and tracking information [88].

Tau Identification

To separate τ-leptons from quark- or gluon-initiated jets, dedicated boosted decision trees (BDTs) are trained for 1-prong and 3-prong τ-leptons, based on several track- and topo-cluster related variables, such as the fraction of the transverse energy deposited in the calorimeter in a certain region around the τ-candidate, or the sum of transverse energy deposited in all cells of the topo-clusters associated with the τ-candidate. For the BDT training, simulated $Z/\gamma^* \to \tau\tau$ events were used for signal, while dijet events were included as background. Based on the BDT output scores, three working points with different τ-selection efficiencies are defined:

- BDT WP Loose: At the loose working point, 1-prong τ-leptons are selected at an efficiency of 60% and 3-prong τ-leptons at 50%.
- BDT WP Medium: The selection efficiency of τ-leptons at the medium working point is 55% for 1-prong τ-leptons and 40% for 3-prong τ-leptons.
- BDT WP Tight: 1-prong τ-leptons have a selection efficiency of 45% and 3-prong τ-leptons 30% at the tight working point.

The reduced efficiencies for the medium and tight working point also come with an increased jet background rejection. Furthermore, the efficiencies are independent of the transverse momenta of the τ-leptons and (as all the BDT input variables are pileup corrected) also do not depend on the number of interactions per bunch crossing. The BDT-based identification of τ-leptons was used in ATLAS analysis until shortly after the end of run 2. The successor of this method is based on a recurrent neural network (RNN), which comes with a largely improved jet rejection power. The input variables for the training of the RNN are very similar to those of the BDT. Dijet events were also used as a background sample for training, while the signal was taken from simulated $\gamma^* \to \tau\tau$ events. Four efficiency working points are defined based on the RNN output score:

- RNN WP Very loose: This working point shows a 1-prong (3-prong) τ-selection efficiency of 95% (95%).
- RNN WP Loose: The loose working points selects 1-prong τ-leptons at an efficiency of 85% and 3-prong τ-leptons with 75%.
- RNN WP Medium: For the medium working point the selection efficiency is 75% for 1-prong and 60% for 3-prong τ-leptons.

- RNN WP Tight: At the tight working point, 60% of the 1-prong τ-leptons and 45% of the 3-prong τ-leptons are selected.

Applying the defined working points in analyses might yield different results for Monte Carlo simulations and data. To account for these differences, scale factors are derived by evaluating the respective efficiencies in data and Monte Carlo simulation via tag-and-probe measurements. The scale factors are applied to the simulated events afterwards [88, 89].

Another type of object that can imitate a τ-signature are electrons. The suppression of electron originated τ-fakes is not covered by the BDTs and RNN that are designed for jet rejection, but it is targeted by two approaches. One exploits the electron likelihood identification (eLLH) (cf Sec. 4.3.1) and performs an overlap removal between the reconstructed electron and the τ-candidate. Thereby, a τ-lepton with only a single track and within $\Delta R < 0.4$ of a reconstructed electron is rejected if the electron fulfils at least the very loose working point based on the electron likelihood identification. The resulting efficiency for hadronic τ-leptons is 95%. The other approach again relies on a dedicated boosted decision tree, separating electrons from hadronically decaying τ-leptons (eBDT). Two working points, referred to as loose (with 95% selection efficiency for τ-leptons) and medium (with an efficiency of 85%). Similar to the jet-τ discriminators, scale factors are calculated from efficiencies measured in tag-and-probe measurements to account for different behaviour of the working points in events from data and simulated events [88, 90].

4.3.5. Missing Transverse Energy

As the partons involved in the collisions only have momentum in z-direction, the sum of all resulting particles' momenta in the transverse plane should also be 0. If it differs from that value, this is an indication of particles created in the collision being undetected. This imbalance is referred to as missing transverse momentum. In the Standard Model of particle physics this effect occurs only for neutrinos, but it is also a commonly expected signature of many theories beyond the Standard Model, including Supersymmetry, where e.g. the lightest neutralino would cause missing transverse momentum.

Missing transverse energy or missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{p}_{\mathrm{T}}^{\mathrm{miss}}|$) is calculated from two major components. First, all the transverse momenta of the reconstructed "hard" objects, such as reconstructed electrons, muons, photons, τ-leptons and jets, are summed up. As the reconstruction algorithms usually are run independently, some signatures in the detector can be part of multiple reconstructed objects. To avoid double counting, an overlap removal is performed before calculating the missing transverse energy. After passing some quality selection criteria, electrons are preferred over other objects in case of shared tracks or topo-clusters, followed by photons, τ-leptons and jets. Muons are not part of this overlap removal as they are predominantly identified from signals in the muon spectrometer and the inner detector. There is, however, a dedicated scheme to resolve overlap between muons and jets. The second part of the missing transverse momentum calculation is the so-called soft-term which takes account of all tracks in the inner detector that origin from the hard-scatter vertex[2] but are not associated with one of the reconstructed hard objects. The derivation of missing transverse momentum can

[2]The hard-scatter vertex is the primary vertex with the largest sum of transverse momenta of tracks associated with it. The other reconstructed primary vertices in an event are due to pileup.

be summarized in the following equation:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = \underbrace{-\sum_{e}\vec{p}_{\mathrm{T}}^{\,e} - \sum_{\gamma}\vec{p}_{\mathrm{T}}^{\,\gamma} - \sum_{\tau}\vec{p}_{\mathrm{T}}^{\,\tau} - \sum_{\mu}\vec{p}_{\mathrm{T}}^{\,\mu} - \sum_{\mathrm{jets}}\vec{p}_{\mathrm{T}}^{\,\mathrm{jet}}}_{\text{hard term}} \underbrace{- \sum_{\text{unused tracks}}\vec{p}_{\mathrm{T}}^{\,\mathrm{track}}}_{\text{soft term}}. \tag{4.11}$$

In addition, the azimuthal angle of the missing transverse momentum is determined as:

$$\phi^{\mathrm{miss}} = \tan^{-1}\left(\frac{E_y^{\mathrm{miss}}}{E_x^{\mathrm{miss}}}\right), \tag{4.12}$$

with E_x^{miss} and E_y^{miss} being the vector components of the missing transverse momentum in x- and y-direction, respectively [91].

4.3.6. Overlap Removal

As already mentioned in the context of missing transverse energy reconstruction (cf Sec. 4.3.5), the algorithms for object reconstruction are mostly run independently from each other and have access to the same detector information. Therefore, it can occur that two different objects are reconstructed from the same tracks or topo-clusters. Consequently, a thorough treatment of the overlaps is of utmost importance. The overlap removal strategy used for the presented analyses is summarized in Tab. 4.3.6. In general the overlap removal consists of consecutive steps and relies mostly on a requirement on the distance in ΔR between the two objects and a predefined priority of objects.

Step	Removed object	Kept object	Condition
1	τ	electron	$\Delta R < 0.2$
2	τ	muon	$\Delta R < 0.2$
3	electron	muon	shared track in the ID
4	jet	electron	$\Delta R < 0.2$
5	electron	jet	$\Delta R < 0.4$
6	jet	muon	number of tracks < 3 and (ghost-associated or $\Delta R < 0.2$)
7	muon	jet	$\Delta R < 0.4$
8	jet	τ	$\Delta R < 0.2$

Table 4.1.: Steps of the overlap removal procedure to avoid ambiguities of the reconstructed objects. In step 5, a "ghost-association" of jet and muon is mentioned. For this, the jet reconstruction algorithm is run again with the muon added with infinitesimal momentum as "ghost" particle. If in the jet reconstruction the muon-ghost is included inside the jet, the jet and the muon are called ghost-associated. The same method is used amongst others for the elimination of muon and jet overlap in the calculation of missing transverse momentum. [91]

4.4. Variable Definitions Based on Reconstructed Objects

In the following sections, the most important variables for the presented analyses will be listed and explained.

4.4.1. Object Multiplicities and Charges

The number of a specific type of object in an event is a crucial information when selecting for a certain final state. Therefore, the number of τ-leptons, the number of light leptons (e,μ) and the number of jets are key elements of each selection. If exactly 0 instances of a certain object type are required, this is referred to as veto. The object multiplicities can also apply only to a specific selection of objects, such as e.g. the number of τ-leptons passing the medium BDT working point ($N(\tau_{\text{medium}})$). Additionally, the charges of the reconstructed particles can play a role, e.g. in requiring oppositely signed electric charge for two objects, which is referred to as "OS" in contrast to "SS" for events with two objects carrying charge with the same sign.

4.4.2. Angular Separation Variables

The concepts of pseudorapidity and the ATLAS coordinate system convention have been introduced in Sec. 3.2.1. The differences in pseudorapidity $\Delta\eta$ as well as the distance in the azimuthal angle $\Delta\phi$ between two objects may differ for physical processes and therefore they can be used as discriminating variables in analyses. Another important use case of these two variables is in the calculation of ΔR between two reconstructed objects, as defined in Eqn. 3.5. Furthermore, there is the possibility to calculate the separation in the azimuthal angle between an object, e.g. a τ-lepton, and the missing transverse energy, for which also the azimuthal information is available. The notation for such a variable would be e.g. $\Delta\phi(\tau, E_{\text{T}}^{\text{miss}})$.

4.4.3. Kinematic Variables

Most kinematic variables depend in some way on the transverse momenta of reconstructed objects and/or the missing transverse energy.

Effective Mass

The effective mass is a measure of the "hardness" of the objects appearing in an event. In the case of the two presented analyses, it is defined as the sum of the absolute values of the tau momenta in the event and the missing transverse energy:

$$m_{\text{eff}} = \sum_{\tau} |\vec{p}_{\text{T}}(\tau)| + E_{\text{T}}^{\text{miss}}. \tag{4.13}$$

Invariant Mass

The Lorentz-invariant mass of a decaying particle (C) is calculated from the four-vectors of its (fully reconstructed) decay products (A and B):

$$p_C^2 = \big(p_A + p_B\big)^2 = m_{\mathrm{inv},C}^2 \tag{4.14}$$

The four-vectors of the decay products A and B can be represented in terms of their transverse momenta $|\vec{p}_{\mathrm{T},A}|, |\vec{p}_{\mathrm{T},B}|$, pseudorapidities η_A, η_B and azimuthal angles ϕ_A, ϕ_B:

$$p_{A/B} = |\vec{p}_{\mathrm{T},A/B}| \cdot \begin{pmatrix} \cosh\eta_{A/B} \\ \cos\phi_{A/B} \\ \sin\phi_{A/B} \\ \sinh\eta_{A/B} \end{pmatrix} \tag{4.15}$$

In the center-of-mass frame of the decaying particle ($\phi_A = -\phi_B = \phi$ and $\eta_A = -\eta_B = \eta$) and given that the decay products A and B are relativistic, i.e. $m_A, m_B \approx 0$, the invariant mass is given as follows:

$$m_{\mathrm{inv},C}^2 = 2 \cdot |\vec{p}_{\mathrm{T},A}| \cdot |\vec{p}_{\mathrm{T},B}| \cdot \big(\cosh(2\eta) - \cos(2\phi)\big). \tag{4.16}$$

The invariant mass is especially useful e.g. to suppress processes containing a Z-boson which decays e.g. into two leptons. In this case, a cut on a certain value of the invariant mass of the two leptons (or a window between two values of the invariant mass) close to the Z-mass is often referred to as Z-veto.

Transverse Mass

Usually in measurements, the mass of a decaying particle is inferred by looking at the invariant mass spectrum. However, this is only possible in the case that all decay products can be fully reconstructed. If one decay product is not reconstructable, e.g. a neutrino from a leptonic W-boson decay, the so-called transverse mass can be used. Taking the example of a W-boson decaying into a τ-lepton and a neutrino, it is defined as follows:

$$m_{\mathrm{T}}^2(\tau) = 2 \cdot E_{\mathrm{T}}^{\mathrm{miss}} \cdot p_{\mathrm{T}}(\tau) \cdot \big(1 - \cos(\Delta\phi(\tau, E_{\mathrm{T}}^{\mathrm{miss}})\big). \tag{4.17}$$

As the neutrino is only "visible" as missing transverse momentum in the detector, its impact can only be estimated in the transverse plane. The transverse mass therefore serves as a projection of the decaying particle's total energy that is carried away by the decay products to the transverse plane. Its maximum, as a consequence, is therefore the mass of the decaying particle and the distribution shows a cut-off at this value. In the discussed example this corresponds to:

$$m_{\mathrm{T}}^2(\tau) \leq m_W^2, \tag{4.18}$$

and therefore plays a large role, e.g. in selecting or rejecting processes containing a W-boson. [92] The sum of transverse masses, in the case of multiple objects in an event, can also be constructed.

Stransverse Mass

The transverse mass tries to "reconstruct" masses of decaying particles if one of the decay products contributes to missing transverse energy. But this is not feasible anymore in scenarios with two particles decaying semi-invisibly. An example for this is the signal process that is targeted in this thesis, where two staus decay into a τ-lepton and a lightest neutralino each, which cannot be detected. The so-called stransverse mass or m_{T2} is a way to recover some information on the decaying particle's mass in case of a pair production and subsequent semi-invisible decay. Suppose, the missing transverse momentum is distributed in some way to the two lightest neutralinos a and b in the case of direct stau pair production and decay:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = \vec{p}_{\mathrm{T}}(\tilde{\chi}^0_{1,a}) + \vec{p}_{\mathrm{T}}(\tilde{\chi}^0_{1,b}). \tag{4.19}$$

Then, in analogy to Eqn. 4.18, the stau mass would obey the following relation, with τ_a and τ_b originating from the same decay as $\tilde{\chi}^0_{1,a}$ and $\tilde{\chi}^0_{1,b}$, respectively:

$$m^2_{\tilde{\tau}} \geq \max\left\{m^2_{\mathrm{T}}(\tau_a, \tilde{\chi}^0_{1,a}), m^2_{\mathrm{T}}(\tau_b, \tilde{\chi}^0_{1,b})\right\}. \tag{4.20}$$

The splitting of the missing transverse momentum to the two invisible particles is not known but it is "guessed" by minimization, giving rise to the variable m_{T2}:

$$m^2_{\mathrm{T2}}(\tau_a, \tau_b) = \min_{\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = \vec{p}_{\mathrm{T}}(\tilde{\chi}^0_{1,a}) + \vec{p}_{\mathrm{T}}(\tilde{\chi}^0_{1,b})} \left[\max\left\{m^2_{\mathrm{T}}(\tau_a, \tilde{\chi}^0_{1,a}), m^2_{\mathrm{T}}(\tau_b, \tilde{\chi}^0_{1,b})\right\}\right]. \tag{4.21}$$

Therefore, by construction, the connection of the stransverse mass to the invariant mass of the decaying particle is as follows:

$$m^2_{\mathrm{T2}}(\tau_a, \tau_b) \leq m^2_{\tilde{\tau}}. \tag{4.22}$$

Similar to the transverse mass, the distribution of the stransverse mass for pair-produced, semi-invisibly decaying particles shows a cut-off at the original particle's mass which makes it one of the most important variables for searches for new physics at collider experiments in general [92].

5. General Aspects of the Search for Stau Pair Production

5.1. General Analysis Strategy

The ultimate goal of the presented searches is to either find evidence for the production of stau pairs in the data or, in the case of a lack of events additional to the amount compatible with the expected Standard Model background, to interpret the result in terms of a minimum non-Standard Model cross section that can be excluded. More specifically, the limits on the cross sections are also interpreted as limits for certain stau and neutralino masses in the simplified model context. To achieve this goal, a region of phase-space has to be identified with an accumulation of stau pair production events (signal) and as low contamination from Standard Model processes (background) as possible. These regions, referred to as signal regions, can be derived by different means of optimization, as is further elaborated in Sec. 6.2 and 7.3. Ultimately, the simulated events are fitted to data in these regions, after what is called unblinding. This means that the signal regions are first designed based entirely on simulated events and only as soon as the analysis and its strategy are fixed, the events from data are taken into account. This has the advantage of avoiding any sort of unconscious bias with respect to possible statistical fluctuations in data. Before performing the fit, however, the simulation of the considered background is checked in dedicated regions enriched with the respective background, which are called validation regions. Those regions can be unblinded right away, because it is ensured that there is no sizeable contribution from signal events. In case of mismodelling that corresponds to a scaling issue, i.e. there are no large slopes in the ratio of data to simulation in the kinematic distributions of concern, control regions can be defined, in which the respective background is normalized during the fit.

There are two distinct iterations of the search for direct stau production at the ATLAS detector with data taken during run 2 of the LHC. The first was published in [51] and is presented in Chapter 6. My personal involvement includes the optimization of the signal region definitions in Sec. 6.2 (together with Chenzheng Zhu) and the construction of the validation regions in Sec. 6.3.3 (in cooperation with Huajie Cheng), as well as the determination of the acceptance and efficiency values for each of the signal models in the signal regions in Chapter 6.7. The second iteration of the analysis proposes improvements compared to the first one and is outlined in Sec. 7. I was personally responsible for all of the presented parts of the second search for stau pair production.

5.2. Statistical Analysis

5.2.1. Statistical Tests

Both the optimization process and the interpretation of the analysis results rely on the evaluation of statistical tests. In the following, the used concepts are briefly introduced.

Given some collection of experimental data $\vec{x}$ with n observed values $(x_1, ..., x_n)$ one can only draw conclusions from in the context of hypotheses H. Usually, there is a null-hypothesis H_0 in contrast to some alternative hypotheses $H_1, H_2, ...$ In the case of a counting experiment as presented in Chapter 6 and Chapter 7 with the purpose of finding a signal in addition to the Standard Model background, the null-hypothesis could e.g. be the number of events that is expected only from background ($n = n_\mathrm{b}$), which is also referred to as background-only hypothesis. In contrast, the signal+background hypothesis is the number of events expected in the presence of a signal ($n = n_s + n_b$). The natural way of interpreting the obtained data is to quantify the degree of agreement with the considered hypotheses. For this, a so-called test statistic is constructed, which depends on the measured data $\vec{x}$. Each of the hypotheses H will yield a different probability density function (p.d.f.) $g(t|H_0)$, $g(t|H_1)$ etc, based on which a decision for the acceptance or rejection of a hypothesis can be made. There are different kinds of test statistics, one of the most intuitive being the so-called p-value. The p-value is defined as the probability of an experimental result or a more "extreme" one, given a certain hypothesis is true. In the case of the counting experiment, this corresponds to the probability of observing n_obs or more events in the experiment under the assumption of e.g. the signal+background hypothesis:

$$p = P(n \geq n_\mathrm{obs}) = \sum_{n=n_\mathrm{obs}}^{\infty} f(n|n_s, n_b). \tag{5.1}$$

In particle physics the p-value is very often converted into the so-called "significance" Z:

$$Z = \Phi^{-1}(1 - p), \tag{5.2}$$

with Φ^{-1} being the inverse of the cumulative standard Gaussian distribution. A Gaussian distributed variable that is found to be Z standard deviations above the mean value has an upper tail probability equal to the p-value. Per convention, in order to exclude a signal+background hypothesis the threshold for the p-value is chosen as 0.05, which corresponds to a significance of 1.64. [70, 93, 94] The significance is also used as a figure of merit for the optimization of the signal regions in Sec. 6.2 and Sec. 7.3

In an experiment, in which for each of the N events the value of a variable x is measured, the collected data can be represented as a histogram $\vec{n} = (n_1, .., n_N)$. The expected value in the i^th bin is then given by

$$E[n_i] = \mu s_i + b_i. \tag{5.3}$$

The parameter μ is referred to as the signal strength, with $\mu = 0$ being equivalent to the background-only hypothesis. s_i and b_i are the mean number of entries in the i^th bin originating from signal and background, respectively. Their values depend on a set of nuisance parameters $\boldsymbol{\theta}$, which can be constrained by constructing another histogram with a dedicated control sample. This sample should only be populated by background and is often used to provide information on the background normalization. From the obtained histograms one can then construct the likelihood function $L(\mu, \boldsymbol{\theta})$ as the product of the

probabilities for all bins, including the bins of the control histogram. A possible test statistic is then the ratio of the likelihoods:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}. \tag{5.4}$$

The numerator represents the maximized likelihood with respect to $\boldsymbol{\theta}$, with the resulting value being $\hat{\hat{\boldsymbol{\theta}}}$, while the signal strength is kept fixed at a chosen value. In the denominator, $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ both maximize the likelihood function. The likelihood ratio will take values between 0 and 1 with larger values corresponding to a better agreement of the data with the hypothesis of the chosen signal strength. An equivalent choice for a test statistic is to take the logarithm of λ:

$$t_\mu = -2 \ln \lambda(\mu). \tag{5.5}$$

Higher values of t_μ correspond to a decrease in compatibility between the data and the hypothesis of signal strength μ. In the case of attempting to set upper limits on the signal strength the slightly changed test statistic is used:

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}. \tag{5.6}$$

This definition makes sure that data with $\hat{\mu} > \mu$ is not regarded as less compatible with μ and is therefore set to 0 in such cases. The agreement of data, represented by an observed $q_{\mu,\text{obs}}$ and the hypothesized μ is further characterized by calculating a p-value:

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu|\mu) dq_\mu, \tag{5.7}$$

with $f(q_\mu|\mu)$ being the p.d.f. of q_μ for a given hypothesized μ, which must be approximated as shown in [93] or calculated with Monte Carlo toy experiments. [70, 93]

To avoid a wrong exclusion of hypothesis in cases, in which the background-only as well as the signal+background hypothesis are disfavoured in a lack of sensitivity, e.g. due to a downward fluctuation in the data, the so-called CL_s value [95] is used:

$$CL_s = \frac{p_\mu}{p_{\mu=0}}. \tag{5.8}$$

Exclusion of the hypothesis μ can be claimed if the CL_s value is smaller than 0.05 per convention, corresponding to a confidence level of 95%.

5.2.2. Fit Types

The fits are performed using the HistFitter analysis framework [96]. All uncertainty sources are taken into account in the form of Gaussian distributed nuisance parameters in the fit, with their width corresponding to the respective uncertainty's size. The number of events in the regions where the fit is performed are treated as Poisson-distributed variables with their mean values being the expected number of events in these regions. The likelihood function that is maximized in the fit is composed as the product of the probability density functions of all the considered regions. There are three different kinds of fits that are performed:

- **Background-only fit**: The background-only fit does not consider any signal contribution and only uses the control regions for the fit, while the signal regions are treated as validation regions. The free parameters in the fit are the normalization factors for the backgrounds that are derived in the control regions as well as the uncertainties.
- **Exclusion fit**: The exclusion fit is also often referred to as model-dependent limit fit. The direct stau signal is taken into account in both the signal and the control regions, which are also included in the fit. The signal contribution is parameterized by a freely floating normalization factor, the already mentioned signal strength. Employing a series of toy experiments or using an approximation method, an upper limit on the signal strength is calculated for each signal model by evaluating the CL_s value in dependence on the signal strength. The upper limit is then given by the value of the signal strength at which the CL_s value drops to 0.05. A signal model (defined by the masses of the stau and the lightest neutralino) is considered as rejected if the upper limit on the signal strength is smaller than 1.
- **Discovery fit**: In the discovery fit (or model-independent limit fit) the one-sided probability $p_0 = p(s = 0)$ under the hypothesis of the absence of signal is calculated that the observed or a larger amount of data events is found. The smaller this value, the clearer are the indications for the presence of a non-Standard Model signal. The value for p_0 is determined for all signal regions separately. The control regions are also included in the fit but under the assumption that they are free of any signal.

5.3. Simulated Processes

Standard Model processes except for the production of multiple jets in the absence of other reconstructed objects are estimated primarily with Monte Carlo simulations and are summarized in classes of backgrounds, as introduced in the following. The details on the employed simulation machinery can also be found here [51].

Z+jets and W+jets

All processes involving a Z-boson (W-boson) produced in the hard scattering in association with an arbitrary number of jets are summarized as Z+jets (W+jets). The generator and parton shower used for the Monte Carlo simulation is provided by Sherpa 2.2.1 [97–99]. Comix [100] and OPENLOOPS [101, 102] were used to calculate the matrix elements at next-to-leading order (NLO) for up to two additional partons and at leading order (LO) for four additional partons, respectively, before the merging with the Sherpa parton shower. The parton density function NNPDF3.0NNLO [71] was employed. The cross sections used for the normalization of the Z+jets and W+jets processes were calculated up to next-to-next-to-leading order (NNLO).

Top

Processes including the production of a number of top quarks are referred to as Top background. This involves single top production and top quark pair production ($t\bar{t}$) as

well as the associated production of top quarks and a heavy boson ($t\bar{t}+V$, $t\bar{t}+H$) and the production of three or four top quarks (multi-top). For single top production and top quark pair production the PowHeg-Box v2 [103–106] generator is used. The matrix elements are calculated at NLO. The NNPDF2.3LO [72] is used as the parton density function set. The parton shower was simulated by Pythia 8.186 [107]. The cross sections are calculated to next-to-next-to-leading order for $t\bar{t}$ and to next-to-leading order for single top production. Events for multi-top production, $t\bar{t}+V$ and $t\bar{t}+H$ are generated with MadGraph5_aMCNLO [108] with the matrix elements calculated at NLO. Pythia 8.186 is providing the parton shower for all simulations of the Top background.

Multiboson

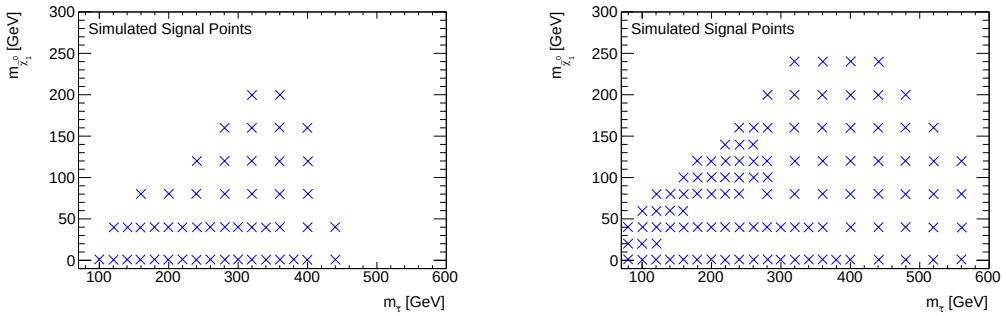
Events with more than one heavy vector boson are grouped as the multiboson background. It includes the production of heavy vector boson pairs (diboson: WW, WZ and ZZ) as well as the production of three vector bosons (triboson). The Sherpa 2.2.1. and 2.2.2 [97, 98] packages are used for event generation and parton showering with NLO accuracy. Again, the NNPDF3.0NNLO set is used for the parton density function.

Higgs

The production of a single Higgs boson (via vector boson fusion as well as via gluon-gluon fusion) and the associated production of a heavy vector boson and a Higgs boson are referred to as Higgs background. For the first, PowHeg-Box v22 was used as event generator and for the latter, the event generation was provided by Pythia 8.186. The parton shower was added with Pythia 8.186 in both cases. The associated production of a Higgs boson with a $t\bar{t}$ pair is simulated using MadGraph5_aMCNLO. The production cross sections are taken from [109].

Direct Stau

To probe the analysis for the supersymmetric target process, a grid of simulations has been created, representing different points in the plane spanned by the masses of the stau (80-560 GeV) and the lightest neutralino (0-240 GeV). The masses of $\tilde{\tau}_L$ and $\tilde{\tau}_R$ are assumed to be degenerate. The masses of squarks, gluinos, charginos and the heavier neutralinos are set to very large value (~ 100 TeV) such that they are not interfering with the target process. The events were generated with MadGraph_aMC@NLO [108] with the parton shower added through Pythia 8.186. The matrix elements were calculated with up to two additional partons and at tree level. As parton density function the NNPDF2.3LO set was used. A detailed list of used mass points and the corresponding cross sections is given in the appendix Sec. A. The cross sections have been derived with the Resummino package [40–44] at next-to-leading order. The signal grid is further illustrated in Fig. 5.1. To kinematically forbid the production of supersymmetric particles other than staus, all chargino and neutralino masses were set to a value of 2.5 TeV.



(a) First Direct Stau Analysis (b) Second Direct Stau Analysis

Figure 5.1.: Each of the blue crosses represents one of the used signal models for the first and second search for stau pair production in dependence of the stau and lightest neutralino masses. In the second search, the amount of studied signal models was increased to cover more of the possible phase space, based on the results of the first stau search.

5.4. Possible Trigger Strategies

Triggers are an important tool to pre-select events with the final state in consideration. As for two hadronically decaying τ-leptons and missing transverse energy there are two kinds of high-level triggers that can be considered, the so-called asymmetric ditau-trigger and the ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger. Both of them are designed to select events with two online reconstructed τ-leptons. The actual implementation of these triggers in terms of threshold values can vary throughout the data taking periods but the general selection criteria stay the same. The asymmetric ditau-trigger is called "asymmetric" because it selects events with at least two "online" τ-leptons[1] with a sizeable difference in the p_{T} thresholds for the leading and second-leading τ-lepton[2]. In general the asymmetric ditau trigger's p_{T}-thresholds are very high which can be beneficial in terms of background suppression but it might also cut away signal events with too low transverse τ-momenta. In contrast, the ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger has lower thresholds on the transverse momenta (while also selecting event with at least two τ-leptons) but in addition it also sets a requirement on the missing transverse energy in an event.

Because triggers usually are not fully efficient at their threshold value on the online object (online threshold), there is a regime called turn-on curve before the trigger reaches its maximum efficiency at some value larger than the online threshold. This marks the beginning of the so-called plateau region in which the efficiency of the trigger stays constant with respect to the object property that is triggered on, e.g. its transverse momentum. While the plateau-region is usually well modelled in simulation as the needed corrections can be applied as a flat factor, this does not necessarily hold for the turn-on region, where p_{T}-dependent mismodelling can occur. To avoid this mismodelling, usually the selection

[1] At HLT level, reconstruction algorithms are used to identify objects with the outcome being referred to as "online" objects. However, the reached efficiency is usually only a fraction of the efficiency quoted for the so-called offline reconstruction algorithms outlined in Sec. 4.3.

[2] The term "leading" object refers to the reconstructed particle with the largest transverse momentum in an event and is denoted as e.g. τ_1 in the case of τ-leptons. The second-leading object is accordingly the reconstructed particle with the second largest transverse momentum in an event, e.g. τ_2.

Trigger	Object Property	Year	Online Threshold [GeV]	Plateau Cut [GeV]
asymmetric ditau-trigger	$p_{\mathrm{T}}(\tau_1)$	2015-2018	80	95
	$p_{\mathrm{T}}(\tau_2)$	2015-2017	50	65
		2018	60	75
ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger	$p_{\mathrm{T}}(\tau_1)$	2015-2017	35	50
		2018	60	75
	$p_{\mathrm{T}}(\tau_2)$	2015-2018	25	40
	$E_{\mathrm{T}}^{\mathrm{miss}}$	2015-2018	50	150
single muon-trigger	$p_{\mathrm{T}}(\mu)$	2015	20	22
		2016-2018	26	28

Table 5.1.: Online and offline thresholds for ditau- and single muon-triggers.

is chosen such that a cut on the offline reconstructed object is placed at the value at which the transition of the turn-on region to the plateau region happens. These cuts are therefore referred to as plateau cuts. In the case of the ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger, the efficiency plateau starts at 150 GeV for an online threshold at 50 GeV. Hence, despite the lower p_{T} threshold for the reconstructed τ-leptons, the ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger introduces already a very tight selection.

Furthermore, the online reconstructed objects are associated with the offline objects in a step called trigger-matching by calculating their distance ΔR. This ensures that the trigger decision is based on the same objects that are used after offline reconstruction. For example, the trigger could in principle select an event due to only one of the two leading τ-leptons and a third one that fails some quality criteria in the offline selection. This can introduce ambiguities e.g. in the application of trigger scale factors. Therefore trigger matching is employed and a third (baseline, cf. Sec. 6.1 and Sec. 6.2.1) τ-lepton is vetoed.

The detailed properties of the asymmetric ditau- and the ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-triggers are listed in Tab. 5.1. For the definition of control regions for the production of a W-boson in association with jets, an event selection with one hadronically decaying τ-lepton and an isolated muon is used (cf. Sec. 6.3.2, 7.4.1). For this purpose, a so-called single muon-trigger is used, which only uses the transverse muon momentum for selection, irrespective of the τ-lepton. Its specifications are also listed in Tab. 5.1.

6. First Search for Direct Stau Production with the Full Run 2 Dataset

6.1. Object Definitions

In the following, the specific definitions of the used reconstructed objects in this analysis are explained, based on the description of the reconstruction and identification methods in Sec. 4.3. In general, objects are split into two categories: "Baseline" and "signal". The baseline objects fulfil looser selection criteria than the signal objects.

Tau Leptons

Baseline τ-leptons are chosen from the set of τ-candidates before any identification working point is applied. Only τ-leptons with a transverse momentum above 20 GeV and within a pseudorapidity range of $|\eta| < 2.5$, excluding the calorimeter transition region $1.37 < |\eta| < 1.52$ are considered. Furthermore, either one (1-prong) or three (3-prong) tracks have to be associated to the baseline τ-leptons, with a total charge of ± 1. Instead of using a specific working point for identification, the baseline τ-leptons have to yield a BDT score of at least 0.01 (for the jet rejection BDT). Signal τ-leptons have to pass all the requirements for baseline τ-leptons and in addition satisfy the medium working point with respect to the jet rejection BDT as well as for the electron rejection BDT.

Electrons

The baseline electrons reconstructed as explained in Sec. 4.3.1 have to pass the loose likelihood identification working point and yield a transverse impact parameter of $|z_0 \sin\theta| < 0.5\,\text{mm}$. To fulfil signal criteria, the electrons furthermore should pass the tight likelihood identification working point and the longitudinal impact parameter has to satisfy $|\frac{d_0}{\sigma(d_0)}| < 5$. For signal electrons, requirements are set as well on the isolation, i.e. they should pass the Gradient isolation working point. At transverse momenta above 200 GeV the HighPtCaloOnly isolation working point is applied.

Muons

Muon candidates with a transverse momentum above 14 GeV and within a pseudorapidity range of $|\eta| < 2.7$ are selected as baseline muons if they also exhibit a transverse impact parameter $|z_0 \sin\theta| < 0.5\,\text{mm}$. In addition, signal muons need to show a longitudinal impact parameter significance of $|\frac{d_0}{\sigma(d_0)}| < 3$. Furthermore, signal muons have to be isolated according to the FCLoose isolation working point.

Jets

Jets are selected from the collection of EMTopo jet candidates at a minimum transverse momentum of 20 GeV and within a range of $|\eta| < 2.8$. A "jet cleaning" procedure is performed to ensure that no jets from sources other than the collision are selected. A jet vertex tagger at a medium working point is used to avoid selecting jets from secondary proton-proton interactions. To identify jets originating from a b-quark the MV2 algorithm is used at an efficiency of 77%.

6.2. Signal Regions

6.2.1. Preselection

The first step to identifying regions which are suitable to isolate the direct stau signal from most of the Standard Model background is to apply selections motivated by the signature of the targeted final state: Two hadronically decaying τ-leptons with oppositely signed electric charge (OS) in the presence of missing transverse energy. Furthermore, lots of Standard Model events can already be eliminated by placing requirements determined by specific signatures of those backgrounds, e.g. the number of b-tagged jets should be set to 0 to suppress Top events (b-veto). Similarly, events containing a Z- or a Higgs boson are rejected by placing a cut on the invariant mass of the two τ-leptons at $> 120\,\text{GeV}$ (Z/H-veto). The collection of these selections together with the trigger selection compose what is referred to as the preselection and is given in Tab. 6.1. Signal events with larger stau masses tend to have a larger missing transverse energy. Therefore, the preselection is split into a "lowMass" region at low values of missing transverse energy and a "highMass" region at higher missing transverse energies. For the lowMass region, the asymmetric ditau-trigger as mentioned in Sec. 5.4 is applied, while in the highMass region the ditau+$E_{\text{T}}^{\text{miss}}$-trigger is used. The latter comes with an offline threshold at a missing transverse energy of 150 GeV. This requirement is inverted for the lowMass region, in order to keep the regions orthogonal. Based on the two regions defined by the preselection cuts, the optimization of the signal regions is performed.

Preselection lowMass	Preselection highMass
2 medium (signal) τ-leptons (OS)	
veto 3$^{\text{rd}}$ baseline τ-lepton	
baseline light lepton veto	
b-tag veto	
Z/H-veto ($m_{\text{inv}}(\tau_1, \tau_2) > 120\,\text{GeV}$)	
asymmetric ditau-trigger	ditau+$E_{\text{T}}^{\text{miss}}$-trigger
$E_{\text{T}}^{\text{miss}} < 150\,\text{GeV}$	$E_{\text{T}}^{\text{miss}} \geq 150\,\text{GeV}$

Table 6.1.: Preselection requirements of first direct stau search. The trigger requirements are understood including the offline thresholds listed in Tab. 5.1.

6.2.2. Optimization and Results

A scan of different combinations of cut variables and cut values is performed and the significance Z is evaluated for each of these cut combinations. This approach is commonly referred to as "cut-and-count" method. The cut variables and values that were tried during this procedure are listed in Tab. 6.2. The significance is calculated with the `RooStats::NumberCountingUtils` package in ROOT [110, 111], taking into account a flat uncertainty of 30% to account for systematic uncertainties, which are not considered in detail at this stage of the analysis. The combination with the maximum significance is chosen if it fulfils the following requirements to ensure sufficient statistics in the signal regions:

- The statistical uncertainty on the total background yield should be smaller than 30%.
- There should be at least one (weighted) background event in total left after all cuts are applied.

One signal region is optimized in the lower $E_{\mathrm{T}}^{\mathrm{miss}}$-region (SR-lowMass) and one for higher values of $E_{\mathrm{T}}^{\mathrm{miss}}$ (SR-highMass). The resulting definition of the signal regions is given in Tab. 6.3 and the respective event yields for all background classes and one of the signal benchmark points are listed in Tab. 6.4.

Variable	Cut Values	
	SR-lowMass	SR-highMass
τ-identification	2 medium τ-leptons, 1 medium +1 tight τ-lepton, 2 tight τ-leptons	
$\mathrm{E_T^{miss}} \geq$	30, 50, 60, 75, 80, 90, 100 GeV	150, 160, 170, 180, 200 GeV
$\mathrm{m_{T2}}(\tau_1, \tau_2) \geq$	40, 50, 60, 70, 80, 90, 100 GeV	
$\mathrm{m_T}(\tau_1) + \mathrm{m_T}(\tau_2) \geq$	200, 250, 300, 350, 400, 450, 500 GeV	
$\Delta\mathrm{R}(\tau_1, \tau_2) \leq$	2.4, 2.6, 2.8, 3.0, 3.2, 6	
$\lvert\Delta\phi(\tau_1, \tau_2)\rvert \geq$	0.4, 0.5, 0.6, 0.8, 1, 1.2	
$m_{\mathrm{inv}}(\tau_1, \tau_2) \geq$	120, 130, 140, 150 GeV	
$\mathrm{p_T}(\tau_1) \geq$	95, 100, 110, 120, 130, 140, 150 GeV	50, 55, 60, 65, 70, 80, 90, 100, 120 GeV
$\mathrm{p_T}(\tau_2) \geq$	60, 70, 80, 90, 100 GeV	40, 45, 50, 55, 60, 70, 80, 90, 100 GeV

Table 6.2.: Cut values scanned during SR optimization. All possible permutations are evaluated for their significance. The cut combination with the largest significance is selected, if the statistical background uncertainty is below 30% and at least one weighted background event is found in this region.

SR-lowMass	SR-highMass
2 tight τ-leptons (OS)	2 medium τ-leptons (OS), ≥ 1 tight τ-lepton
veto 3$^{\mathrm{rd}}$ baseline τ-lepton	
baseline light lepton veto	
b-tag veto	
Z/H-veto ($m_{\mathrm{inv}}(\tau_1, \tau_2) > 120\,\mathrm{GeV}$)	
$\lvert\Delta\phi(\tau_1, \tau_2)\rvert > 0.8$	
$\Delta R(\tau_1, \tau_2) < 3.2$	
$m_{\mathrm{T2}}(\tau_1, \tau_2) > 70\,\mathrm{GeV}$	
asymmetric ditau-trigger	ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger
$75 < E_{\mathrm{T}}^{\mathrm{miss}} < 150\,\mathrm{GeV}$	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 150\,\mathrm{GeV}$

Table 6.3.: Signal region definitions for first direct stau search. The main separation power originates from the cut on m_{T2}.

The chosen cut combinations are further motivated in so-called N-1 plots. These plots show the distributions of signal and background with all cuts applied except for the cut on the variable that is plotted. The behaviour of the significance in dependence of this variable is used to check the validity of the chosen cuts. The most powerful variable in

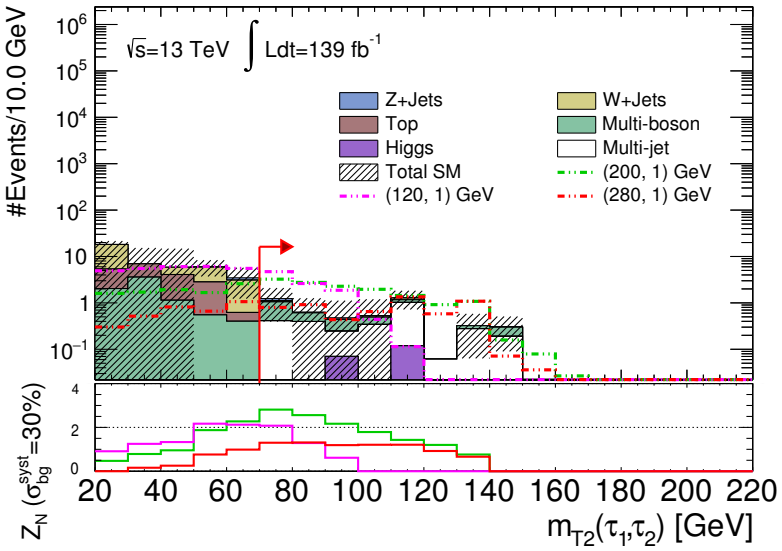
	SR-lowMass	SR-highMass
Z+jets	0.45 ± 0.18	$0.040^{+0.058}_{-0.04}$
W+jets	$0.00^{+0.28}_{-0.00}$	1.9 ± 1.7
Top	0.041 ± 0.041	1.96 ± 0.87
Multiboson	1.37 ± 0.32	2.57 ± 0.42
Higgs	0.0113 ± 0.0071	–
Multijet	2.87 ± 0.36	1.8 ± 1.6
SM total	4.73 ± 0.59	8.2 ± 2.5
$(m_{\tilde{\tau}}, m_{\tilde{\chi}^0_1}) = (200, 1)$ GeV	14.16 ± 0.90	19.8 ± 1.3

Table 6.4.: Event yields in SR-lowMass and SR-highMass. Only the statistical uncertainties are taken into account at this stage.

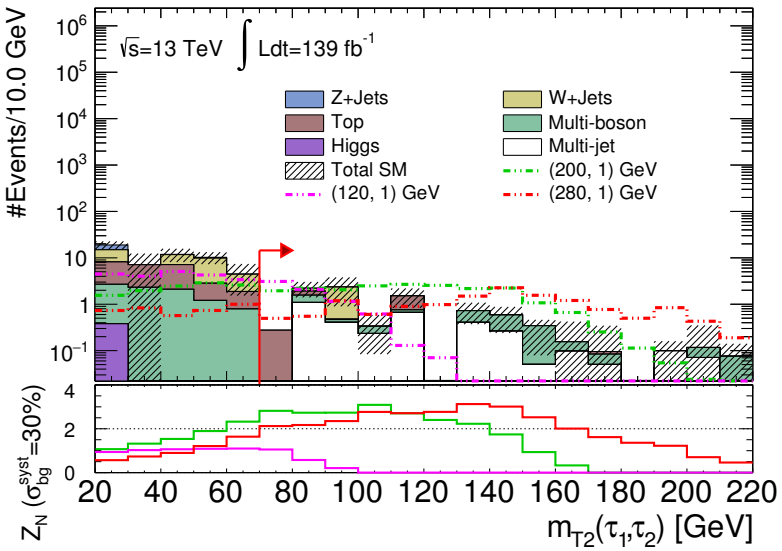
terms of separation of signal and background is m_{T2}. The N-1 plots for m_{T2} are shown in Fig. 6.1.

By evaluating the significance for all available signal models, a first estimate of the exclusion sensitivity can be made, which is shown in Fig. 6.2 for both SR-lowMass and SR-highMass separately and the combination of the two regions. The estimated combined significance is calculated as

$$Z_{\mathrm{combined}} = \sqrt{Z^2_{\text{SR-lowMass}} + Z^2_{\text{SR-highMass}}}. \tag{6.1}$$

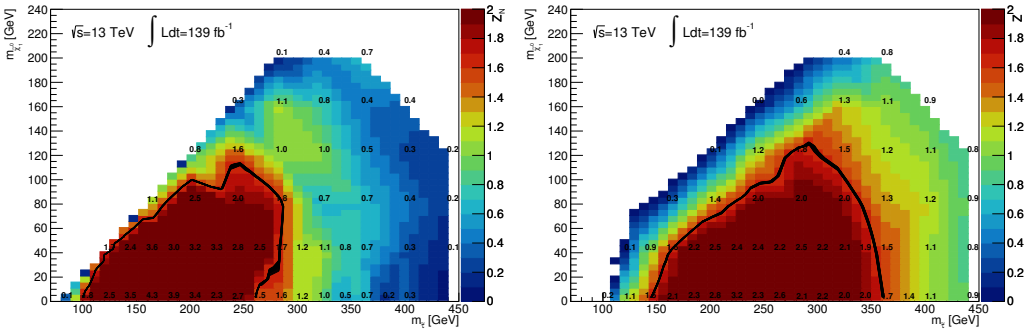


(a) m_{T2} in SR-lowMass



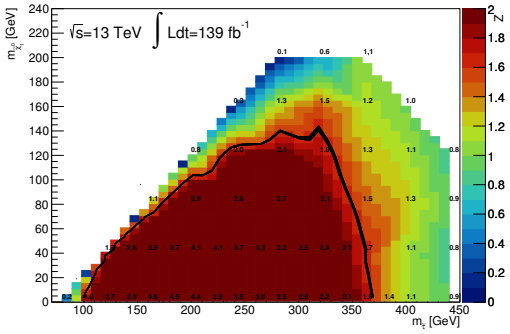
(b) m_{T2} in SR-highMass

Figure 6.1.: N-1 plots of m_{T2} for SR-lowMass and SR-highMass. The red arrow indicates the cut on the m_{T2} variable itself. The coloured lines correspond to the used signal benchmark points, for which the stau and the neutralino masses are given in brackets in the legend. The lower panel shows the significance for a hypothetical lower cut at each of the bins in the histogram.



(a) SR-lowMass

(b) SR-highMass



(c) Combined Significance

Figure 6.2.: Significance per signal model for SR-lowMass, SR-highMass and their combination. As mentioned in Sec. 5.2, expected exclusion sensitivity is reached at a significance value of at least 1.64. This critical value is marked by a black line.

6.3. Background Estimation

In the following, the estimation of the contributions of the separate background categories on top of the Monte Carlo simulations given in Sec. 5.3 will be elaborated in more detail.

6.3.1. Multijet Estimation

Events containing a number of jets originating from strong interaction processes are summarized as multijet background. It contributes to the signal regions with events containing two jets, being wrongly reconstructed as τ-leptons. Because this type of processes cannot be calculated perturbatively, a data-driven method called ABCD-method is employed. The main idea is to design a set of three control regions close to the considered signal region from which the multijet contribution can be extrapolated. As these control regions should not contain much signal events, they can be unblinded and the multijet contribution can be estimated by subtracting the Monte Carlo (MC) generated events for the other Standard Model backgrounds from the data events. The ratio of multijet contribution is determined with two control regions (CR-B and CR-C) that are separated from each other by a certain requirement and from the third control (CR-A) region and the signal region (SR-D) by another selection. This ratio is referred to as transfer factor and can be applied to the estimated multijet contribution in the third control region to obtain an estimate for the multijet contamination in the signal region. The transfer factor is calculated as

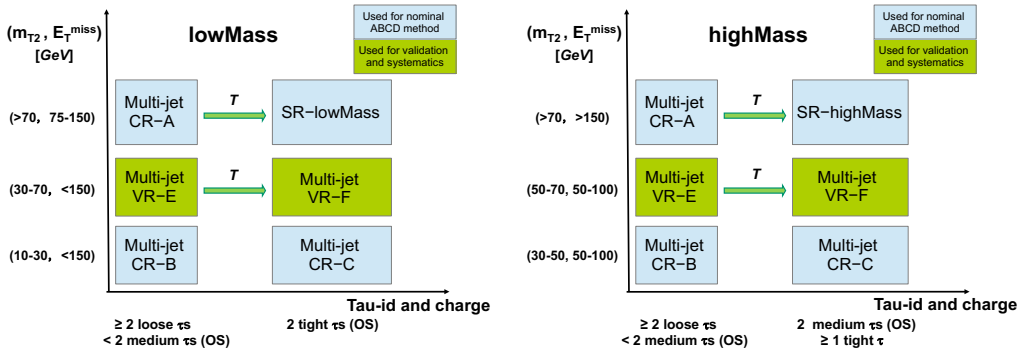
$$T = \frac{\text{Data(C)-MC(C)}}{\text{Data(B)-MC(B)}}. \tag{6.2}$$

To ensure that the transfer factor is also valid for CR-A and SR-D the selection criteria separating CR-A from SR-D and CR-B from CR-C must be independent from the requirement separating CR-B and CR-C from CR-A and SR-D. In this case, the first criterion consists of requirements on the τ-identification working point and the relative sign of the two τ-charges while the latter is composed of cuts on m_{T2} and the missing transverse energy. Additionally, two validation regions, VR-E and VR-F, serve to validate the method and provide information on the effect of small correlations between the previously mentioned requirements. The transfer factor is applied to the difference of data and simulated events in VR-E to estimate the multijet background in VR-F. This estimate is compared to the actual multijet contribution in VR-F by calculating the difference of data to simulated events in this region as well. The resulting deviations later-on enter as systematic uncertainties to the fit.

The ABCD-method as well as the regions definitions are illustrated in Fig. 6.3. More details to the multijet estimation can be found in [51].

6.3.2. W+jets Estimation

Events with one W-boson and a number of jets contribute by one τ-lepton originating from the decay of the W-boson and the other one being a misidentified jet. Besides the Monte Carlo simulation mentioned in Sec. 5.3 two methods are used to improve the reliability of this background's estimation. Firstly, it is normalized in a control region in the fit and secondly the used statistics of simulated events is increased by the τ-promotion method,



(a) ABCD-method for SR-lowMass (b) ABCD-method for SR-highMass

Figure 6.3.: Illustration of ABCD method. The transfer factor T is calculated from the ratio of the multijet contributions in CR-B and CR-C. It can then be applied to the multijet contribution in CR-A (VR-E) to obtain an estimate for the multijet contamination in SR-D (VR-F). [51]

which was specifically developed for this analysis and is briefly described in the following subsection.

The normalization of the W-background performed in the control region (W-CR) is checked in a dedicated validation region (W-VR). Both are enriched in W+jets events and use a selection requiring one signal muon and one medium τ-lepton. The definitions of both regions are given in Tab. 6.5. In addition to the veto of b-tags, it is required that no top-tags appear in an event. The top-tagging is based on the m_{CT} variable introduced in [112].

W-CR	W-VR
single muon-trigger (cf. Sec. 5.4)	
1 medium tau and 1 signal muon (OS)	
$p_{\mathrm{T}}(\tau) > 60\,\mathrm{GeV}$	
$p_{\mathrm{T}}(\mu) > 50\,\mathrm{GeV}$	
$50 < m_{\mathrm{T}}(\mu) < 150\,\mathrm{GeV}$	
$m_{\mathrm{inv}}(\tau,\mu) > 70\,\mathrm{GeV}$	
$E_{\mathrm{T}}^{\mathrm{miss}} > 60\,\mathrm{GeV}$	
$m_{\mathrm{T}}(\tau) + m_{\mathrm{T}}(\mu) > 250\,\mathrm{GeV}$	
$1 < \Delta R(\tau,\mu) < 3.5$	
b-tag veto	
top-tag veto	
$30 < m_{\mathrm{T2}}(\tau,\mu) < 70\,\mathrm{GeV}$	$m_{\mathrm{T2}}(\tau,\mu) > 70\,\mathrm{GeV}$

Table 6.5.: W-CR and W-VR definitions. One medium τ-lepton and one signal muon are required.

Tau Promotion Method

One limiting factor for the sensitivity for direct stau production in the signal regions (cf. Sec. 6.2) is that the statistical uncertainty on the number of W+jets events becomes very large especially for tighter selections. The reason for this is that one of the τ-leptons in a selected W+jets event has to be a so-called fake τ-lepton, i.e. a jet misidentified as a τ-lepton. The tighter the chosen τ-identification working point, the less simulated (unweighted) events will pass the selection, thus leading to an increased statistical uncertainty. The τ-promotion method has been developed to increase the statistics for W+jets by recovering otherwise rejected events with fake τ-leptons that pass only loser or no τ-identification requirements. The events have to be reweighted afterwards to restore the original number of weighted events. The additional weight is calculated through the rate of fake τ-leptons. The information of whether a reconstructed τ-lepton is actually a jet (fake τ-lepton) or is correctly reconstructed (true τ) is determined in a process called truth matching. A τ-lepton is called truth matched if, when accessing the truth information stored for the Monte Carlo simulation (cf. Sec. 4.2.3), the reconstructed τ-lepton lies within a cone of $\Delta R < 0.2$ of a τ-object at truth level. If this requirement is not fulfilled, the τ-lepton is referred to as fake τ-lepton or non-truth matched τ-lepton. Considering e.g. an event with one true τ-lepton (from the W-decay) and otherwise no reconstructed object passing the required τ-identification working point, one randomly chosen fake τ-lepton is "promoted" to a higher τ-identification working point (medium or tight).

The fake rate or fake efficiency that is needed to reweight the additionally selected weights afterwards, is measured in bins of $p_{\mathrm{T}}(\tau)$ and missing transverse energy, as well as for 1-prong and 3-prong τ-leptons separately. It is defined as

$$\epsilon = \frac{\text{number of non-truth matched reconstructed } \tau \text{ passing signal ID}}{\text{number of non-truth matched reconstructed } \tau}, \tag{6.3}$$

with "signal ID" referring to the considered τ-identification working point, i.e. medium or tight. After performing the combinatorics, the additional weighting factor for events with "promoted" τ-leptons is found to be

$$w = \frac{\epsilon}{1-\epsilon} \cdot \frac{\text{number of non-signal non-truth matched } \tau \text{ before promotion}}{\text{number of signal non-truth matched } \tau \text{ after promotion}}. \tag{6.4}$$

The systematic uncertainty introduced by the tau promotion method is estimated by comparing the event yields of the W+jets sample before and after tau promotion in the signal regions when applying all selection criteria sequentially. The largest deviation in the yields of the "nominal" and the "promoted" samples is used as the systematic uncertainty ($\tilde{2}5\%$ for both signal regions).

This method is further described in [113].

6.3.3. Estimation of Other Backgrounds

Besides multijet and W+jets events, a large part of the background processes originate from irreducible backgrounds, which means that they yield the same final state or a final state indistinguishable from the direct stau signal process. Such processes arise from Z+jets or Higgs events, in which the Z-boson or the Higgs boson decays into two τ-leptons, events with two or more vector bosons decaying into τ-leptons and neutrinos

(multiboson) and partly also events with one or more top quarks (although the latter can cause events with fake τ-leptons as well, which would qualify as reducible background). For all these backgrounds, the plain Monte Carlo simulation is taken as an estimate for their contributions in the signal regions without e.g. a normalization in a control region. Nevertheless, the modelling of the simulations is checked in dedicated validation regions, one for each the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ and the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ region, for the background categories of Z+jets, Top and Multiboson. As the contribution of the Higgs boson background is very small in the signal regions, no dedicated validation region is needed for this background. The definitions of the validation regions are listed in Tab. 6.6.

Variable	LowMass			HighMass		
	T-VR	Z-VR	VV-VR	T-VR	Z-VR	VV-VR
	≥ 2 medium τ-leptons (OS), ≥ 1 tight τ-lepton					
	≥ 1 b-tag	b-tag veto		≥ 1 b-tag	b-tag veto	
$m_{\mathrm{inv}}(\tau_1, \tau_2)$	–	$< 70\,\mathrm{GeV}$	$< 110\,\mathrm{GeV}$	–	$< 60\,\mathrm{GeV}$	$< 110\,\mathrm{GeV}$
$\Delta R(\tau_1, \tau_2)$	> 1.2	< 1	–	> 1.2	> 1	–
$m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$	–	–	$> 250\,\mathrm{GeV}$	–	–	$> 200\,\mathrm{GeV}$
$m_{\mathrm{T2}}(\tau_1, \tau_2)$	$> 60\,\mathrm{GeV}$	$< 60\,\mathrm{GeV}$	$> 60\,\mathrm{GeV}$	$> 60\,\mathrm{GeV}$	$< 60\,\mathrm{GeV}$	$> 60\,\mathrm{GeV}$
	asymmetric ditau-trigger $60 < E_{\mathrm{T}}^{\mathrm{miss}} \leq 150\,\mathrm{GeV}$			ditau+$E_{\mathrm{T}}^{\mathrm{miss}}$-trigger $E_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$		

Table 6.6.: Validation region definitions. These regions are used to check the modelling of the backgrounds with which they are enriched. They are, however, not an input to the likelihood fit.

6.4. Systematic Uncertainties

Besides the systematic uncertainties introduced by the ABCD method for the estimation of the multijet background (Sec. 6.3.1) and the τ-promotion method for the improvement of the W+jets estimate (Sec. 6.3.2), there are several other uncertainties that have to be considered. One can distinguish between experimental uncertainties that arise e.g. through the reconstruction of objects and theoretical uncertainties due to the choice of the event generator inputs and the background modelling. All systematic uncertainties enter as nuisance parameters into the fit. The dominant sources of uncertainties are given in Tab. 6.7.

6.4.1. Overview of Experimental Uncertainties

The main sources of experimental uncertainties in the signal regions are caused by the identification of τ-leptons and the tau energy scale, as well as the jet energy scale and resolution. Another contribution is due to resolution and scale of the soft term in the calculation of the missing transverse energy. Other experimental uncertainties, like uncertainties on the energy resolution, energy scale and identification of reconstructed objects other than τ-leptons and jets or pileup reweighting, are also taken into account but are found to be of minor impact to the signal regions (less than 3%).

6.4.2. Overview of Theoretical Uncertainties

The most important theoretical uncertainties are due to the SHERPA event generator setup used for the background categories of Z+jets, W+jets and multiboson and are determined by varying the set of parton density functions according to the PDF4LHC recommendations [114]. Furthermore, variations of the QCD renormalization and factorization scales are taken into account. These uncertainties for W+jets and multiboson processes are the dominant source of theoretical uncertainties and constitute 2-3% on the total background yield for the former and 5-6% for the latter in the signal regions.

Uncertainties for the simulation and modelling of events with a top-quark pair ($t\bar{t}$) arise through the parton shower and the modelling of initial and final state radiation. The parton shower uncertainty is evaluated by comparing simulations with the Pythia 8 parton shower and with the parton shower by Herwig7 [115]. The initial and final state radiation uncertainties are estimated by varying the radiation settings in the simulation process.

For all background processes as well as the signal samples, uncertainties on the production cross sections are taken into account.

	Source of systematic uncertainty	SR-lowMass (%)	SR-highMass (%)
Background	Statistical uncertainty of MC samples	11	21
	Tau identification and energy scale	19	10
	Normalisation uncertainties of the multi-jet background	13	9
	Multijet estimation	6	11
	W+jets theory uncertainty	5	8
	Multiboson theory uncertainty	5	6
	Jet energy scale and resolution	5	8
	$E_{\mathrm{T}}^{\mathrm{miss}}$ soft-term resolution and scale	2	2
	Total	28	33
	Source of systematic uncertainty	SR-lowMass (%)	SR-highMass (%)
	$\mathrm{m}(\tilde{\tau},\tilde{\chi}_1^0)$ GeV	(120, 1)	(280, 1)
Signal	Tau identification and energy scale	29	14
	Statistical uncertainty of MC samples	6	10
	Signal cross section uncertainty	4	6
	Jet energy scale and resolution	3	2
	$E_{\mathrm{T}}^{\mathrm{miss}}$ soft-term resolution and scale	3	< 1
	Total	31	18

Table 6.7.: Dominant relative uncertainties. The upper table shows the uncertainties for the total background yield in the signal regions, while the lower table represents a benchmark signal point with a stau mass of 120 GeV (280 GeV) for SR-lowMass (SR-highMass) and a lightest neutralino mass of 1 GeV. The numbers represent the relative uncertainties after the fit. [51]

6.5. Results

In the following, the results of the three different fit types are presented. There are two normalisation factors that are determined inside control regions: One for the multijet background inside the CR-A from the ABCD method (cf. Sec. 6.3.1) and one for the W+jets background class inside the control region W-CR (cf. Sec. 6.3.2).

6.5.1. Background-Only Fit

The background-only fit is performed as explained in Sec. 5.2.2, with only the number of events inside the control regions being subject to the fit and the signal regions treated as validation regions. The normalisation derived for the multijet background is 1.03 with an uncertainty of ± 0.36. For the W+jets background the normalisation is measured as 0.91 ± 0.12. The resulting number of events and the comparison with observed data events in the validation region is shown in Fig. 6.4. The number of events after the fit in the signal and control regions are listed in Tab. 6.8. In all the validation regions a good agreement of observed and expected events is found. In SR-lowMass more events are observed than expected while in SR-highMass fewer events have been observed than were expected. These deviations are, however, not significant as can be seen from Tab. 6.8 and the m_{T2} distribution for SR-lowMass and SR-highMass after the fit in Fig. 6.5 [51].

SM process	Multijet CR -lowMass	Multijet CR -highMass	W-CR	SR -lowMass	SR -highMass
Multiboson	1.4 ± 0.6	1.9 ± 1.0	63 ± 21	1.4 ± 0.8	2.6 ± 1.4
W+jets	13 ± 4	4^{+7}_{-4}	850 ± 70	1.5 ± 0.7	2.5 ± 1.8
Top quark	2.7 ± 0.9	3.3 ± 1.6	170 ± 40	$0.04^{+0.80}_{-0.04}$	2.0 ± 0.6
Z+jets	$0.25^{+1.43}_{-0.25}$	1.5 ± 0.8	13 ± 7	$0.4^{+0.5}_{-0.4}$	$0.05^{+0.13}_{-0.05}$
Multijet	55 ± 10	16 ± 6	–	2.6 ± 0.7	3.1 ± 1.4
SM total	72 ± 8	27 ± 5	1099 ± 33	6.0 ± 1.7	10.2 ± 3.3
Observed	72	27	1099	10	7

Table 6.8.: Event yields after the background-only fit for the signal and the control regions. The fit is only performed in the control regions, thereby determining the normalization factors for the multijet and W+jets backgrounds. The given uncertainties correspond to the combined statistical and systematic uncertainty on the event yields. The Higgs background was removed from this table due to negligible contributions. Table adapted from [51].

6.5.2. Model Independent Cross Section Limits

By performing a fit in each of the signal regions separately (still using the control regions) the expected and observed p_0-values are calculated and upper limits on the cross section for non-Standard Model processes are derived. Tab. 6.9 shows the resulting values together with the expected event yields of two benchmark signal models for comparison. The background-only hypothesis cannot be rejected given the calculated p_0-values. The upper

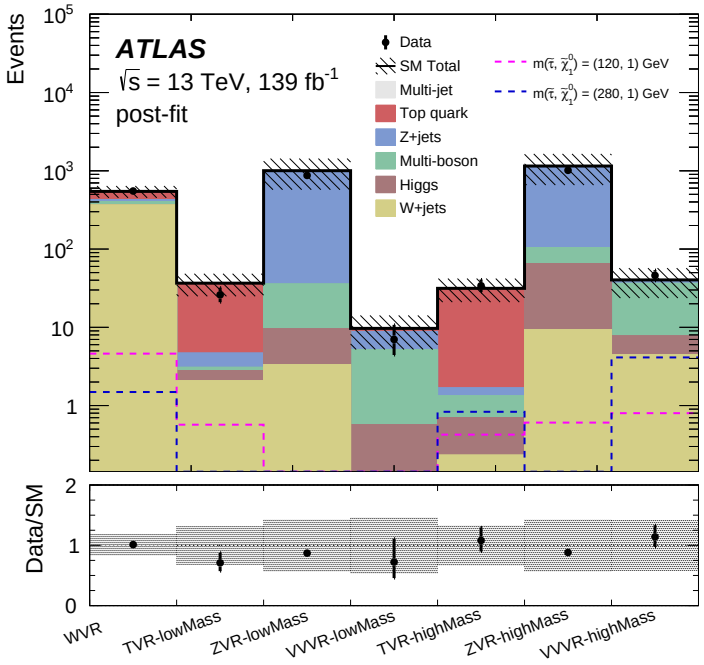
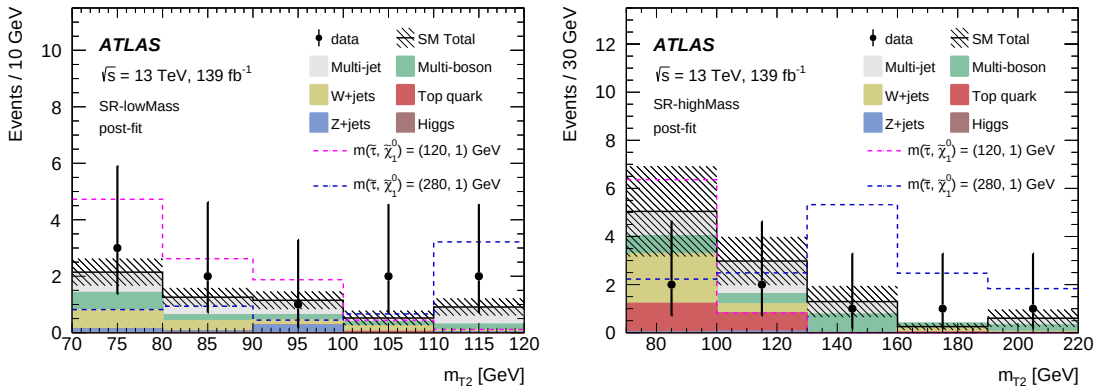


Figure 6.4.: Total number of events in the validation regions after the background-only fit. The normalisation factors determined in the fit have been applied to the the multijet and W+jet backgrounds. The lower panel shows the agreement of data to the estimated Standard Model background. In all of the regions the observed and the expected number of events agrees within the combined statistical and systematic uncertainties. Two direct stau signal models have been added for illustrative reasons with the pink and the blue dashed lines. [51]

limits on the visible cross section σ_{vis} (defined as the product of the production cross section, the reconstruction efficiency and the selection acceptance) are therefore very small, namely 0.08 fb for SR-lowMass and 0.05 fb for SR-highMass [51].



(a) SR-lowMass (b) SR-highMass

Figure 6.5.: Post-fit $m_{\mathrm{T2}}(\tau_1, \tau_2)$ in the signal regions. In general the fitted background expectation agrees well with the observed data. It can be seen that the excess in SR-lowMass is mainly due to one bin in the m_{T2} histogram. In SR-highMass the slight overestimation of events can also be attributed to a single bin in m_{T2}. For comparison, two signal models have been added to the plots as pink and blue dashed lines. [51]

	SR-lowMass	SR-highMass
$m(\tilde{\tau},\tilde{\chi}_1^0) = (120, 1)\,\mathrm{GeV}$	9.8 ± 3.1	7.2 ± 2.2
$m(\tilde{\tau},\tilde{\chi}_1^0) = (280, 1)\,\mathrm{GeV}$	5.9 ± 1.5	14.0 ± 2.5
p_0	0.11	0.50
Expected $\sigma_{\mathrm{vis}}^{95}$ [fb]	$0.055^{+0.025}_{-0.014}$	$0.065^{+0.025}_{-0.019}$
Observed $\sigma_{\mathrm{vis}}^{95}$ [fb]	0.08	0.05

Table 6.9.: Discovery fit results. The p_0-value for each of the signal regions is shown as well as the upper limit on the non-Standard Model cross section. p_0 values larger than 0.5 are truncated at this value. Neither of the p_0-values indicates a disagreement with the background-only hypothesis in the signal regions. The upper limits on the cross sections are including the acceptance of the signal region selections and the reconstruction efficiency and are given at 95% confidence level. [51]

6.6. Interpretation: Model Dependent Limits

Because there is no significant excess of data over the Standard Model expectation observed in the signal regions, the results can be interpreted in terms of exclusion limits for signal models depending on the masses of the stau and the lightest neutralino. These limits are obtained by performing an exclusion fit (or model-dependent limit fit as explained in Sec. 5.2.2). A signal model is excluded when the signal strength μ_{sig} at which the CL_s-value drops below 0.05 (corresponding to a 95% CL) is smaller than 1. The observed (expected) excluded pairs of stau and lightest neutralino masses are beneath the solid red (dashed black) line in Fig. 6.6. The yellow band corresponds to the $\pm 1\sigma$ variations on the expected exclusion limits considering all uncertainties except the theoretical cross section uncertainties for the signal models. The latter are shown separately as the dotted red line around the observed exclusion limit. Fig. 6.7 shows the same interpretation in terms of exclusion limits but only taking into account the pair production of the supersymmetric partner of the left-handed tau lepton [51].

Summarizing the exclusion contours shown in Fig. 6.6 and Fig. 6.7, stau masses can be excluded between 120 GeV and 390 GeV for massless lightest neutralinos in the context of the simplified model under consideration and for the combination of $\tilde{\tau}_L$ and $\tilde{\tau}_R$ pair production. Taking only pair production of $\tilde{\tau}_L$ into account, the excluded stau masses range from 155 GeV to 310 GeV [51].

Exclusion limits could not be derived for $\tilde{\tau}_R$ pair production separately, however. This is mainly due to the lower cross section for $\tilde{\tau}_R$ pair production as compared to $\tilde{\tau}_L$ pair production, despite the acceptance times efficiency being larger for the former due to slight kinematic differences. The according studies will be shown in Sec. 6.7.

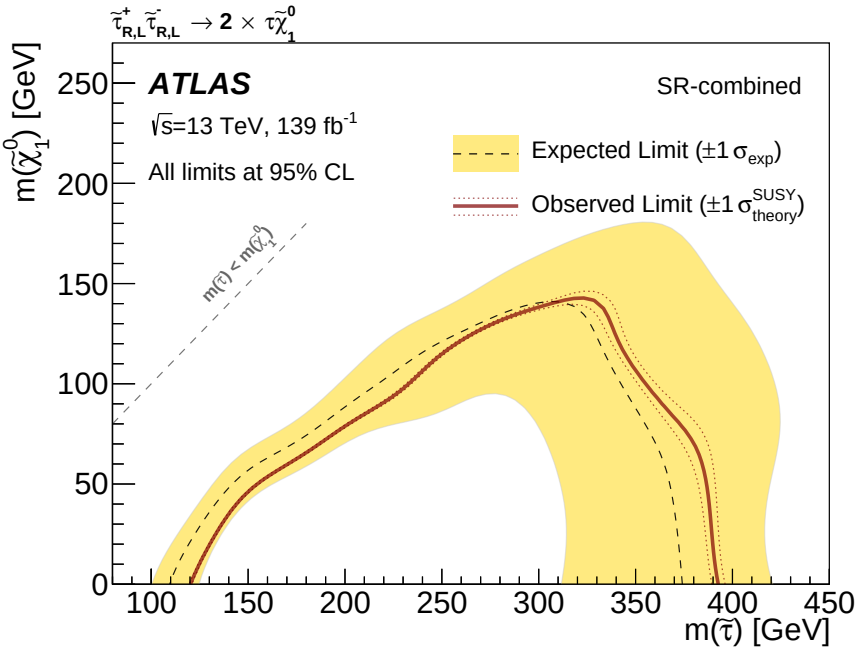


Figure 6.6.: Exclusion contour for $\tilde{\tau}_L$ and $\tilde{\tau}_R$ pair production. All pairs of stau and lightest neutralino masses beneath the solid red line are excluded with 95% CL in the context of the considered simplified model. The dashed black line corresponds to the expected exclusion limit. The observed limit does not extend to as low stau masses as the expected limit due to the small excess observed in the lowMass signal region. Similarly the observed exclusion includes higher stau masses than the expected limits because there are slightly less observed events in the highMass signal region than were expected. The yellow band corresponds to the $\pm 1\sigma$ variations. [51]

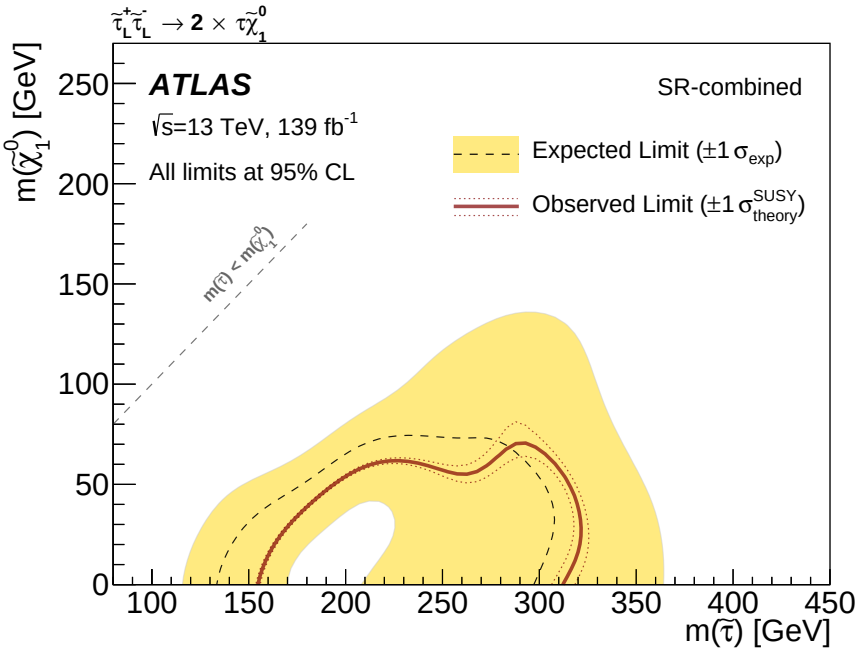


Figure 6.7.: Exclusion plot for $\tilde{\tau}_L$ pair production. The red line shows the observed exclusion limit and the dashed black line the expected limit. Separate exclusion of stau and neutralino mass points for $\tilde{\tau}_R$ pair production could not be achieved. [51]

6.7. Acceptance and Efficiency in the Signal Regions

In the following section, the acceptance and efficiency of the signal regions requirements (cf. Sec. 6.2) are investigated. This is done for the two signal regions taking the pair production of the $\tilde{\tau}_L$ and $\tilde{\tau}_R$ eigenstates into account separately as well as their combination.

6.7.1. Acceptance

The signal acceptance α is calculated as the fraction of events that is selected by the preselection and signal region requirements at truth level:

$$\alpha = \frac{\text{Number of signal events in SR at truth level}}{\sigma \cdot \mathcal{L}}. \tag{6.5}$$

Here, σ is the SUSY production cross section and $\mathcal{L}$ the integrated luminosity ($139\,\text{fb}^{-1}$). The resulting values for each signal model are shown for the highMass and lowMass signal regions in Fig. 6.8 for the combination of both stau eigenstates. For $\tilde{\tau}_L$ and $\tilde{\tau}_R$ pair production separately, the plots can be found in the appendix in Fig. C.1 and Fig. C.2. It can be seen that the acceptance tends to be larger for the pair production of the $\tilde{\tau}_R$ eigenstate than for the $\tilde{\tau}_L$ eigenstate. This is true for both signal regions, although SR-highMass shows in general better signal acceptance than SR-lowMass.

6.7.2. Efficiency

The signal reconstruction efficiency ϵ is given by the ratio of signal events in the signal region at reconstruction level to the signal yield in the signal region at truth level:

$$\epsilon = \frac{\text{Number of signal events in SR at reco level}}{\text{Number of signal events in SR at truth level}}. \tag{6.6}$$

The respective values for different signal mass points are shown in Fig. 6.9 for the highMass signal and lowMass regions for the combined stau eigenstates. The corresponding plots for separate $\tilde{\tau}_L$ and $\tilde{\tau}_R$ pair production can be found in the appendix in Fig. C.3 and Fig. C.4. For SR-highMass the reconstruction efficiency has a tendency to be larger than for SR-lowMass, especially for higher stau masses. However, the lowMass signal region shows large reconstruction efficiency only at a very small stau mass ($< 100\,\text{GeV}$), while the efficiency is rather similar for the rest of the considered models.

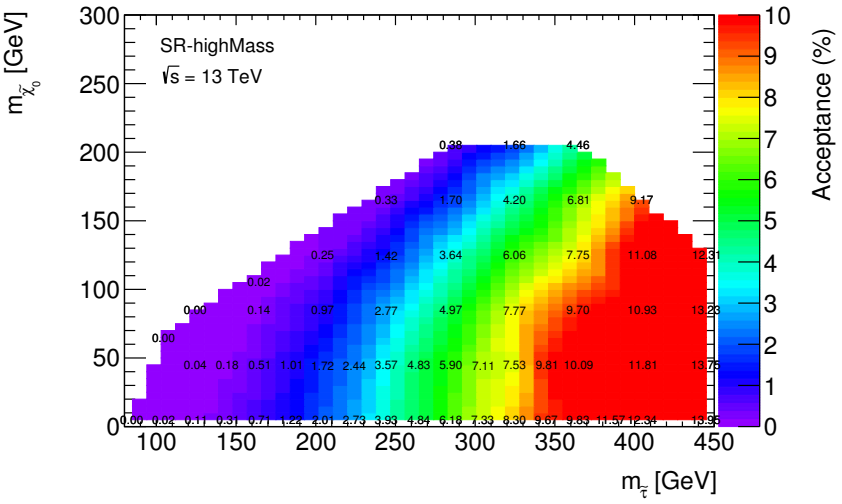
6.7.3. Acceptance×Efficiency

The product of signal acceptance and signal efficiency is calculated as the fraction of reconstructed events in the signal region among the total number of events:

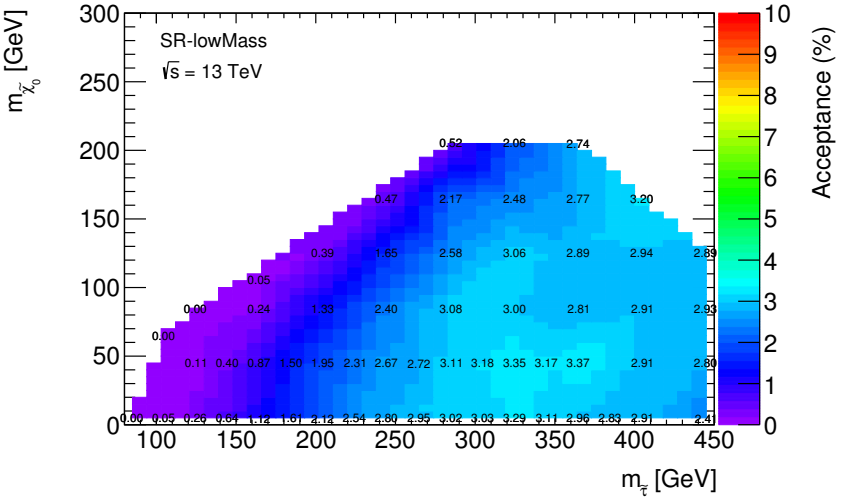
$$\alpha \cdot \epsilon = \frac{\text{Number of signal events in SR at reco level}}{\sigma \cdot \mathcal{L}}. \tag{6.7}$$

The results are shown in Fig. 6.10 for the highMass and the lowMass signal regions for the combination of both stau eigenstates. The plots for the separate eigenstates are given in the appendix in Fig. C.5 and Fig. C.6. Again, in SR-highMass the product of acceptance and efficiency is in general larger than for SR-lowMass, especially for stau masses $> 250\,\text{GeV}$.

Signal models with large stau masses might in general have better acceptance and efficiency values in both signal regions than lower stau masses, but it is still not easy to reach sensitivity for very high stau masses due to the decreasing cross section. Lower stau masses, though not so much favoured by the current signal region selections, have comparatively larger cross sections. An extension of the previously shown sensitivity towards lower stau masses is however necessary as the limits derived by LEP only exclude stau masses up to about 90 GeV. Similarly, although being preferentially selected in the signal regions, it was not possible to achieve sensitivity for signal events from $\tilde{\tau}_R$ production separately due to their much smaller cross section compared to $\tilde{\tau}_L$ pair production.

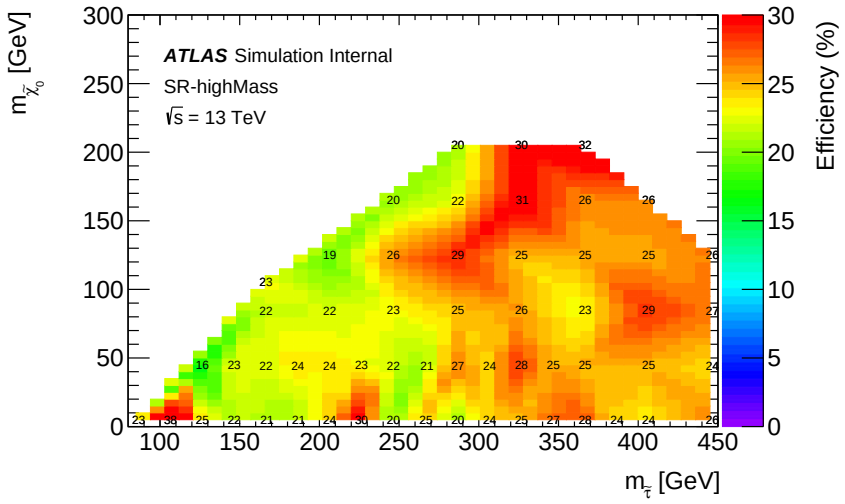


(a) SR-highMass

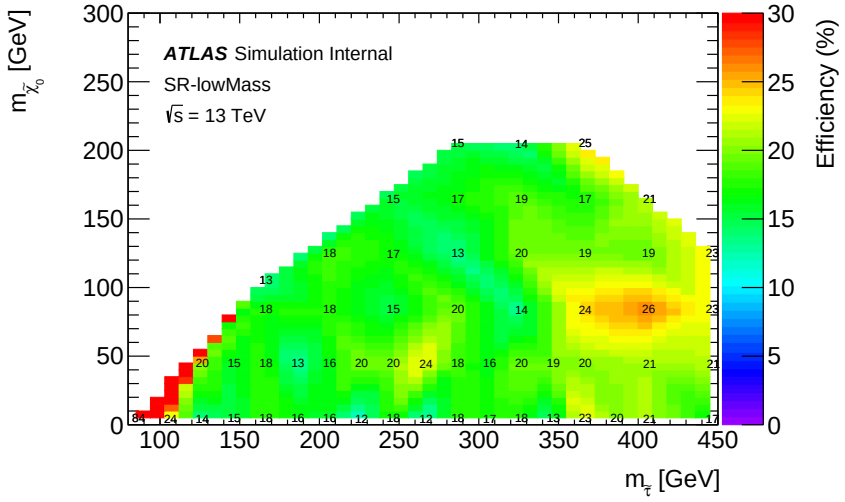


(b) SR-lowMass

Figure 6.8.: Signal acceptance for different stau and neutralino masses in the highMass and lowMass signal regions for the combined stau eigenstates.

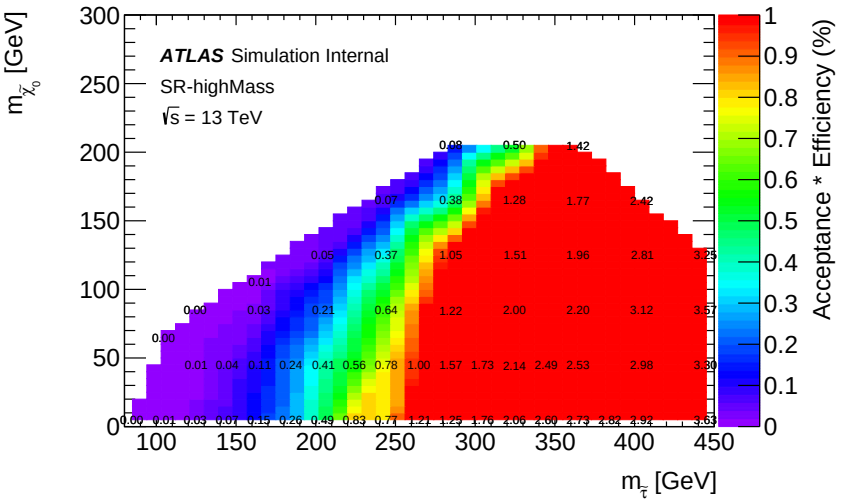


(a) SR-highMass

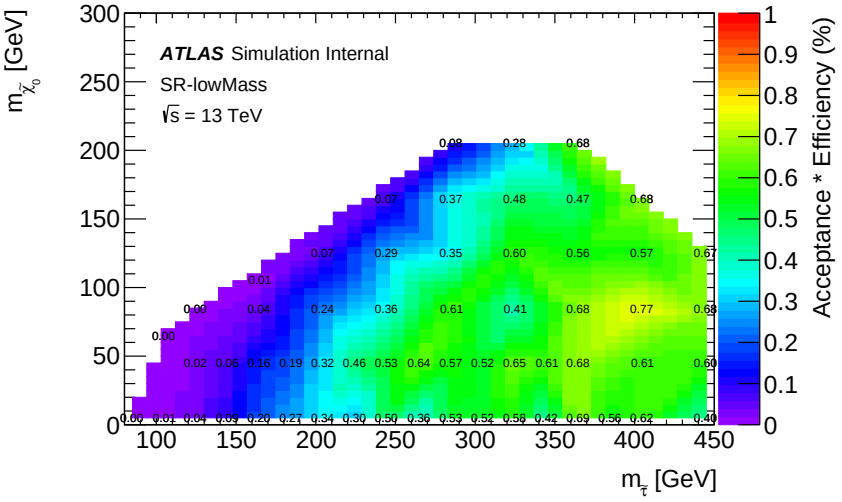


(b) SR-lowMass

Figure 6.9.: Signal efficiency for different stau and neutralino masses in the highMass and lowMass signal regions for the combined stau eigenstates.



(a) SR-highMass



(b) SR-lowMass

Figure 6.10.: Signal acceptance×efficiency for different stau and neutralino masses in the highMass and lowMass signal regions for the combined stau eigenstates.

7. An Improved Search for Direct Stau Production

Although the first analysis of the full dataset taken by ATLAS in run 2 yielded unprecedented sensitivity for several models of direct stau production, extending the exclusion limits on the stau mass to 390 GeV, there is still room for improvement. On the one hand, it was not possible to extend the sensitivity to stau masses as small as the LEP limit ($\approx 90\,\text{GeV}$). The exclusion limit also leaves a large gap to the kinematically forbidden region (at $m_{\tilde{\tau}} > m_{\tilde{\chi}_1^0}$). Furthermore, the studies in Sec. 6.7 are indicating that despite the small cross section for $\tilde{\tau}_R$ pair production, it might be possible to achieve sensitivity for this channel as the product of acceptance and efficiency is in general larger for $\tilde{\tau}_R$ than for $\tilde{\tau}_L$ pair production. Finally, by improving the discrimination of signal and background the analysis could also increase its sensitivity to models with low production cross sections, i.e. even larger stau masses.

The analysis presented in the following is aimed at improving some of the previously mentioned points with respect to the first iteration of the direct stau search.

7.1. Objects Definitions

The definition of objects for this analysis is in most parts identical to that for the first direct stau search (cf. Sec. 6.1). There are, however, two notable differences that are related to the improvements of the jet reconstruction and τ-lepton identification, besides a few minor modifications that will be briefly summarized in the following.

Tau Leptons

The most important difference in terms of object definitions between the two iterations of the analysis is the availability of an improved τ-identification scheme for the second iteration, namely four new working points based on an RNN score for the discrimination of τ-leptons and jets. For τ-leptons at the baseline level, there is only a minimum requirement for the RNN score of 0.05. To pass signal criteria, the τ-leptons have to pass the medium RNN working point.

Electrons

Compared to the first analysis iteration, the isolation working point for signal electrons is changed from the Gradient to the Fixed-Cut Loose working point. For electrons with transverse momenta above 200 GeV the Fixed-Cut HighPtCaloOnly working point is used.

Muons

The definition of muons is identical to that in the first direct stau search with the exception that the minimum transverse momentum required at baseline level was lowered to 3 GeV while for signal muons at least 25 GeV are required.

Jets

Another important change with respect to the first analysis iteration is the usage of PFlow jets instead of EMTopo jets. Furthermore, while in the first direct stau search the MV2 decision tree discriminant for the b-jet identification has been used, the DL1 b-tagger, based on a deep neutral network, is employed in the second iteration. The efficiency working point for the b-tagger is still at 77%.

7.2. Machine Learning and Boosted Decision Trees

The first iteration of the direct stau search relied strongly on the discrimination power of one single kinematic variable, namely m_{T2}. For the second iteration of the analysis new discriminators are constructed based on machine learning techniques that can combine the discrimination power of several variables. Such a new variable would separate signal (the direct stau events) and background (the Standard Model processes) in an optimized way. In data science this kind of challenge is referred to as a classification problem.

7.2.1. Classifiers and Boosting

There are many approaches to handle classification problems with machine learning techniques. Perhaps one of the conceptually most straightforward are decision trees. In principle decision trees work similarly to the cut-and-count method that was presented in Sec. 6.2. A kinematic variable with some discrimination power is selected and an optimized cut is placed. Afterwards the background and signal purity below and above the cut value is checked.[1] If the respective purities in the two regions defined by the cut are below a certain threshold, another variable is selected and a cut value is optimized. This procedure is iterated until the purity threshold is reached or a given maximum of splittings is hit (referred to as maximum depth). Depending on whether a region is pure in signal or in background, it is referred to as signal-like or background-like. Following the decision tree steps, events can be classified depending on whether they are selected into a signal-like or a background-like region. The principle of a decision tree is illustrated in Fig. 7.1.

In general, it is possible to construct powerful discriminators by means of decision trees but they quickly hit similar limitations as a cut-and-count strategy, i.e. low statistics in the final regions (leaf nodes). There are, however, ways to create a strong discriminator by combining an ensemble of "weak learners", where a weak learner is a discriminator that performs only slightly better than random guessing. The principle of (hypothesis) boosting was first introduced by R. Schapire in 1990 [116]. There are different kinds of boosting, e.g. adaptive boosting (AdaBoost) [117, 118] which relies on reweighting individual events after the construction of each decision tree depending on the misclassification of events. This way, wrongly classified events are considered more strongly for the next decision tree. After the desired number of decision trees was created, the final discriminator is constructed as the average over all decision trees.

Another conceptual approach to boosting is to consider this problem as the minimization of a certain loss function $L(y, F(\vec{x}))$, with y being the label of an event (e.g. -1 for background and $+1$ for signal) and $F(\vec{x})$ a function that provides an approximate mapping of the input features (i.e. variables) to the labels y. There are many possible ways to define a loss function, e.g. by the difference of the mapped value and the actual label $|y - F(\vec{x})|$. For this analysis, the negative binomial log-likelihood function has been used:

$$L(y, F(\vec{x})) = \ln\left(1 + e^{-2yF(\vec{x})}\right). \tag{7.1}$$

The minimization of this loss function is performed numerically via the steepest descent (also gradient descent) algorithm. An essential requirement for this is that the loss func-

[1] This is only possible if any information on the nature of the events is available, i.e. if events are "labelled" as signal or background. This type of machine learning is called supervised learning.

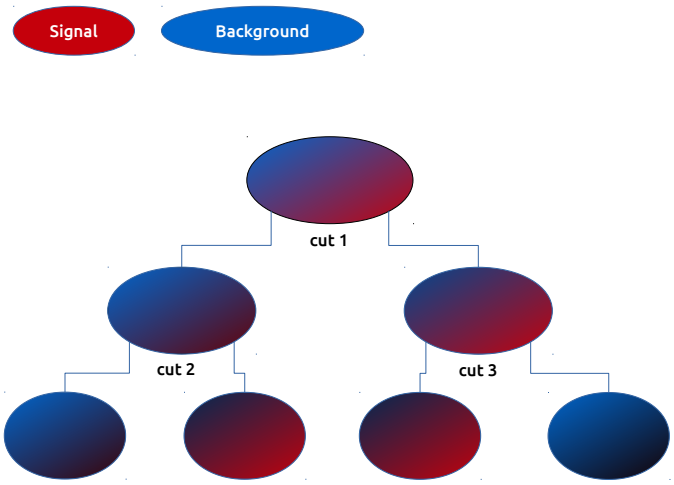


Figure 7.1.: Illustration of the concept of decision trees. The ellipses represent the composition of events in a region before a cut is placed. The red colour stands for signal events while background events are indicated by blue. Cuts are placed iteratively on different variables in order to increase the purity of signal and background in the regions until the maximum number of consecutive cuts or a purity threshold is reached.

tion is differentiable. The gradient of the loss function at the current iteration stage is calculated and a regression tree is fitted to it. The total discriminator is then updated with the fitted regression tree, thereby minimizing the loss function at each step. Actually, the gradient boosting is only applicable to regression problems but can be also used for classification problems by assigning one regression model per class and interpreting their outputs as probability for an event to belong to that class.

Another important parameter often used with the iterative creation of classifiers is the so-called shrinkage or learning rate. It is used to decrease the impact of individual learners on the total classifier to improve the generalization performance and to reduce overtraining [119–121].

7.2.2. Setup of a Boosted Decision Tree

Besides external inputs to a boosted decision tree, like the number of iterations (N(trees)), the maximum depth of the individual trees and the shrinkage (referred to as hyperparameters), there are a few more ingredients to be considered. Concerning the dataset used to train the classifier, the input variables have to be chosen. This is often done by using shape plots to assess the discrimination power added by each potential variable beforehand. Furthermore, it might also be beneficial in some cases to apply an event selection prior to training, e.g. if one particular background (of the total Standard Model background) is clearly dominating a certain region of phase-space to make sure the BDT considers other backgrounds and other regions as well. Finally, the problem of generalization has to be considered as well. As mentioned previously, the shrinkage is used to control the effect

of overtraining, i.e. the training on statistical fluctuations. If applied to different events than to those on which the BDT was trained, an overtrained discriminator would yield a very high misclassification rate, while performing almost perfectly on the training data. Exactly this means of checking overtraining is employed by dividing the input dataset into two subsets: One training dataset and one testing dataset. The training is only done with the former, while the resulting classifier is evaluated on the latter and checked if the performance is sufficiently similar for the two subsets. These comparisons are usually done in terms of plots, e.g. by comparing the output scores directly or the ROC curves[2], as well as quantitatively, e.g. by comparing the misclassification rates. It is also possible to use a Kolmogorov-Smirnov test to determine whether the output score distributions are sufficiently similar. The test result can be interpreted as a probabilistic statement as to whether the hypothesis of the two distributions being the same is to be rejected. A very small Kolmogorov-Smirnov test statistic indicates a good agreement of the two distributions. In this analysis, the Kolmogorov-Smirnov tests are performed with the SciPy library. [122, 123]

All the ingredients of a BDT mentioned previously are in principle subject to optimization. The hyperparameters are optimized by consecutively testing different combinations of hyperparameter values and thereby maximizing the accuracy (correctness of classification) and the ROC value[3]. The sample fractions used for training and testing are determined by balancing the background and the signal samples. For the signal sample, a fixed fraction of 80% (20%) is used for training (testing). As there are typically more background events available than signal events, the fraction of background events used for training is adjusted to approximately match the number of signal training events. The remaining background events are used for testing. As the choice of kinematic variables at each stage of the training is optimized for separation power anyway, in principle every variable can be passed to the BDT. However, with the help of variable rankings it can be determined which variables were selected most often and thus had the largest impact on the training result. [124]

The training and evaluation of classifiers for this analysis has been performed by using functionalities from the scikit-learn machine learning library [125].

[2]Receiver Operating Characteristic: The ROC curve shows the background rejection as a function of the signal acceptance. For a maximal discrimination of signal and background, the shape of the curve would resemble a square. A “discriminator” based on random guessing would yield the diagonal of that square as its ROC curve.

[3]The ROC value is the area under the ROC curve, which is normalized to 1. A large ROC value is usually an indication for good discrimination power.

7.3. Signal Region Optimization and Definition

As the signal models show notably different kinematic behaviour in dependence on the stau mass (and also on the mass difference of stau and lightest neutralino), a similar strategy is used as in the first iteration of the direct stau search: A distinction is made between the "lowMass" and the "highMass" regime (in reference to the stau mass). Therefore, two discriminators in the form of boosted decision trees are required, one for each regime, which are called LowMassBDT and HighMassBDT. The main difference between these two BDTs is that they are trained on different sets of signal models. While the LowMassBDT is trained for stau masses between 80 and 160 GeV, stau masses between 360 and 440 GeV are used for the HighMassBDT training. The input for the background class consists of the Monte Carlo simulations introduced in Sec. 5.3 and a data-driven estimate for the multijet background, which will be explained in more detail in Sec. 7.4.5. Both BDTs are trained with events selected by the asymmetric ditau-trigger. To help avoiding overtraining, the trigger plateau cuts are removed for training to increase the training dataset. The specifications and optimized hyperparameters as well as pre-training selections for the BDTs are listed in Tab. 7.1 and Tab. 7.2. The lists of input variables and their ranking are given in Tab. D.1 and Tab. D.2 in the appendix. The individual discrimination power of the input variables can be checked from shape plots, which are shown in the appendix in Sec. D.2. The linear correlation parameters between each input variable pair can be found in the form of matrices in Sec. D.3. The output scores and ROC curves are shown in Fig. 7.2. The upper plots of Fig. 7.2 show the normalized distribution of the output scores for signal (red) and background (blue) and for both testing (filled area) and training (lines) datasets. In addition, the resulting Kolmogorov-Smirnov test statistic is given for the comparison of the testing and training scores for both signal and background. In all cases, the Kolmogorov-Smirnov test yields a very small value which indicates that the distributions are compatible and suffer from negligible overtraining. This can also be observed from the ROC curves, which are plotted for both the testing and training dataset as well. The area under the curve is not only very large for both BDTs (thus indicating a good separation power), it is also in good agreement for the training and testing dataset. The histograms in Fig. 7.3 show again the BDT output scores but evaluated on the full dataset and stacked for all backgrounds and a few signal benchmark models. Some of the bins at lower BDT scores, where the signal contamination is very small, are unblinded to check the agreement of data and background expectation. The lower panel of these plots shows the evolution of the significance with rising BDT output score for the signal benchmark models.

Based on these output scores, the signal regions for the fit setup can be defined. For the highMass signal region (SR-highMass) a cut on a HighMassBDT score of 0.9 has been optimized. This signal region will be used for both the exclusion and the discovery fit. The LowMassBDT shows a very smooth rise in significance with increasing score for lower stau mass points, which is not as steep and localized in the highest score bins as for the HighMassBDT (for higher stau masses). Especially signal models with lower stau masses seem to benefit not only from very high output scores but also from upper intermediate output scores above 0.5. For this reason, a shape fit of the LowMassBDT score at > 0.5 is performed in four equidistant bins to derive model-dependent limits. Since it is not possible to perform a discovery fit in more than one bin, an "alternative" SR-lowMass region is designed for obtaining model-independent limits. For this region, a cut on the LowMassBDT score is optimized for significance and placed at > 0.75. In addition, the

lowMass and highMass signal regions are separated by a cut on the missing transverse energy at 150 GeV. The definitions for the signal regions can be found in Tab. 7.3 and Tab. 7.4. The expected number of background events as well as the expected yield for some signal benchmark models are given in Tab. 7.5 for the SR-lowMass shape fit and in Tab. 7.6 for the SR-lowMass discovery region and SR-highMass. The expected significances for all regions are shown in Fig. 7.4.

Parameter	Setting
Model	GradientBoost
N(trees)	500
Maximum Depth	3
Shrinkage	0.1
Loss function	"deviance" (cf. scikit-learn documentation [125])
Signal	$80 \leq m_{\tilde{\tau}} \leq 160\,\text{GeV}$ $1 < m_{\tilde{\chi}_1^0} < 100\,\text{GeV}$ $\tilde{\tau}_\text{R}\tilde{\tau}_\text{R}$ and $\tilde{\tau}_\text{L}\tilde{\tau}_\text{L}$
Background	Z+jets, W+jets, Multiboson, Multijet
Selection (pre-training)	asymmetric ditau-trigger (without plateau cuts) at least two medium τ-leptons OS $\Delta R(\tau_1, \tau_2) < 4$ $m_{\text{T2}}(\tau_1, \tau_2) > 30\,\text{GeV}$ $m_{\text{vis}}(\tau_1, \tau_2) > 80\,\text{GeV}$

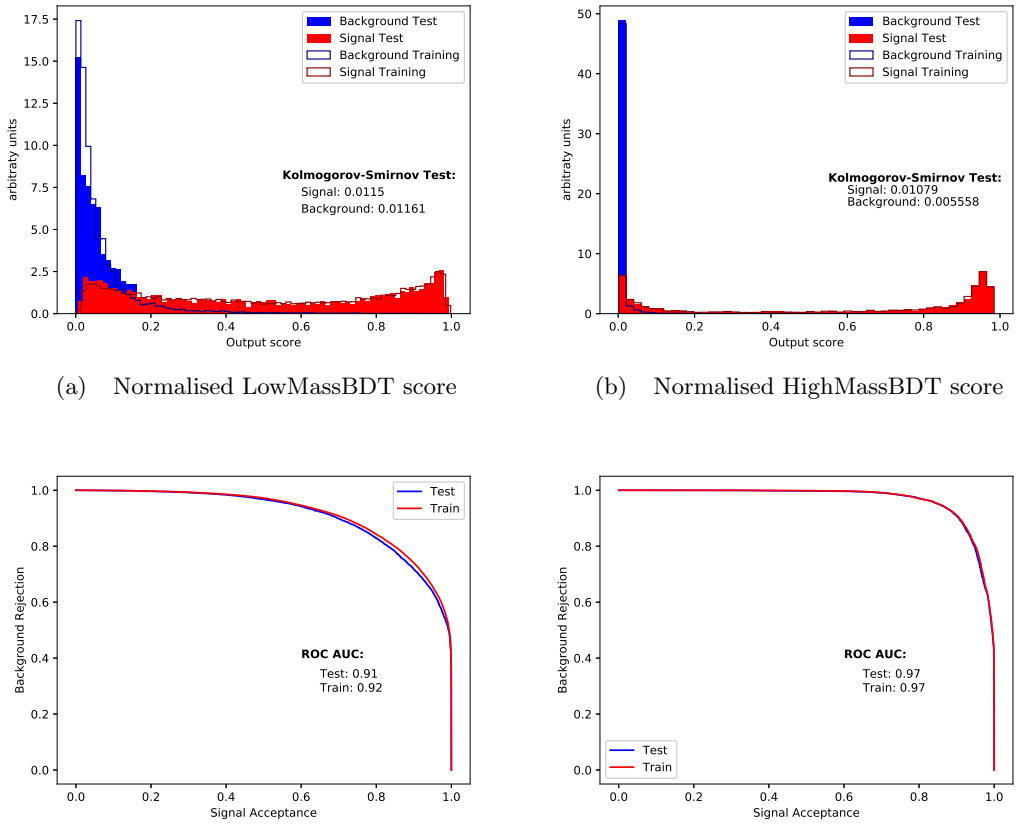
Table 7.1.: Training setup for the Direct Stau LowMassBDT. Only signal mass points with $80 \leq m_{\tilde{\tau}} \leq 160\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} \leq 100\,\text{GeV}$ have been used for training. The signal samples are split randomly into subsets of 80% training events and 20% testing events. The number of background events for training is chosen such that it is in the same order of magnitude as the number of training events for the signal sample. The remaining background events are used for testing. For the estimation of the multijet background a simplified ABCD method has been employed (cf. Sec. 7.4.5). The plateau cuts are not applied for training in order to increase the number of training events. This makes the training procedure less prone to overtraining. The trigger plateau cuts are, however, applied after training for the SR-, CR- and VR-definitions.

Parameter	Setting
Model	GradientBoost
N(trees)	500
Maximum Depth	3
Shrinkage	0.01
Loss function	"deviance"
Signal	$360 \leq m_{\tilde{\tau}} \leq 440\,\text{GeV}$ $1 < m_{\tilde{\chi}_1^0} < 40\,\text{GeV}$ $\tilde{\tau}_\text{R}\tilde{\tau}_\text{R}$ and $\tilde{\tau}_\text{L}\tilde{\tau}_\text{L}$
Background	Z+jets, W+jets, Multiboson, Top, Multijet
Selection (pre-training)	asymmetric ditau-trigger (without plateau cuts) at least two medium τ-leptons OS $m_\text{vis}(\tau_1, \tau_2) > 80\,\text{GeV}$

Table 7.2.: Training setup for the Direct Stau HighMassBDT. Only signal mass points with $360 \leq m_{\tilde{\tau}} \leq 440\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} \leq 40\,\text{GeV}$ have been used for training. The signal samples are split randomly into subsets of 80% training events and 20% testing events. The number of background events for training is chosen such that is is in the same order of magnitude as the number of training events for the signal sample. The remaining background events are used for testing. For the estimation of the multijet background a simplified ABCD method has been employed (cf. Sec. 7.4.5).

SR highMass	SR lowMass (exclusion)			
2 medium (signal) τ, OS asymmetric ditau trigger 3$^\text{rd}$ baseline τ veto light lepton veto b-veto $m_{\text{T}2}(\tau_1, \tau_2) > 30\,\text{GeV}$ $p_\text{T}(\tau_1) > 95\,\text{GeV}$ $p_\text{T}(\tau_2) > 65\,\text{GeV}$ $m_\text{vis}(\tau_1, \tau_2) > 80\,\text{GeV}$				
–	$\Delta R(\tau_1, \tau_2) < 4$			
$E_\text{T}^\text{miss} > 150\,\text{GeV}$	$E_\text{T}^\text{miss} \leq 150\,\text{GeV}$			
hmBDT > 0.9	$0.5 <$ lmBDT < 0.625	$0.625 <$ lmBDT < 0.75	$0.75 <$ lmBDT < 0.875	lmBDT > 0.875

Table 7.3.: Signal region definitions for the exclusion Fit. "hmBDT" and "lmBDT" are short for HighMassBDT and LowMassBDT, respectively.



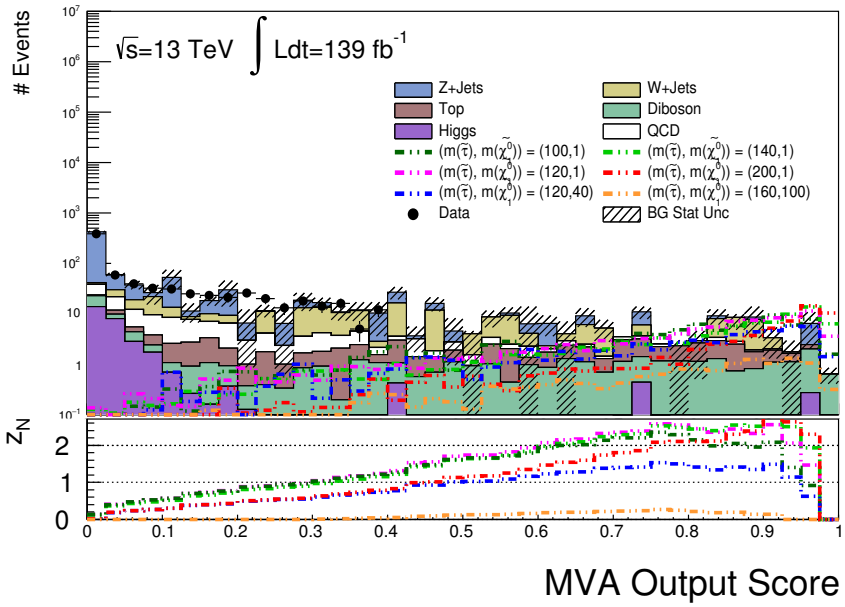
(a) Normalised LowMassBDT score

(b) Normalised HighMassBDT score

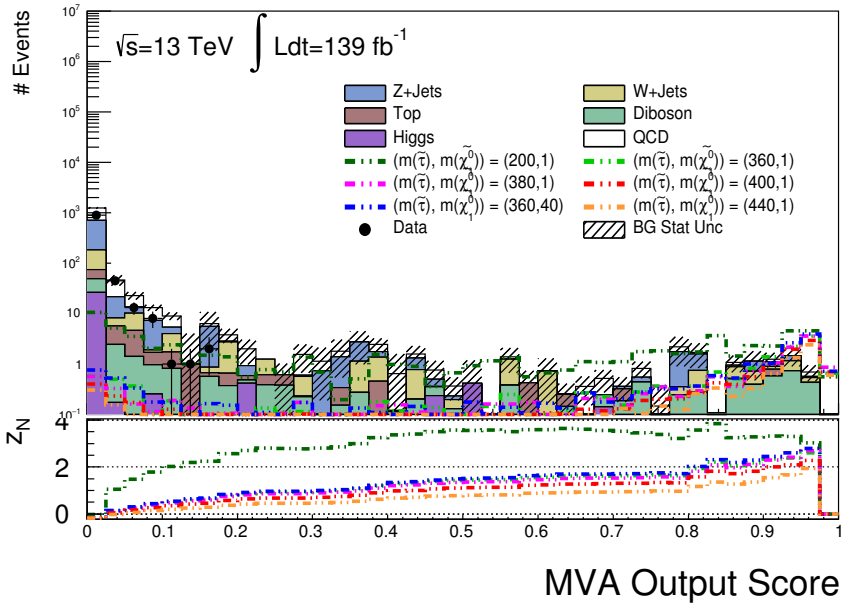
(c) ROC curve LowMassBDT

(d) ROC curve HighMassBDT

Figure 7.2.: BDT output scores and ROC curves. The upper plots show the normalised output scores for the LowMassBDT and the HighMassBDT for the training (line) and the testing dataset (filled area) as well as for signal (red) and background (blue). A Kolmogorov-Smirnov test is performed to determine the compatibility of the training and testing distributions. It returns a KS test statistic, which indicates good agreement at very small values. The two plots in the lower row show the ROC curves for both BDTs again for the training (red line) as well as for the testing (blue line) dataset. Both curves show good agreement, which can also be seen from the areas under the curve calculated for the training and testing set. Both the LowMassBDT with an area under the curve of 0.9 and the HighMassBDT with 0.97 show furthermore a good separation power of signal and background.



(a) LowMass BDT score



(b) HighMass BDT score

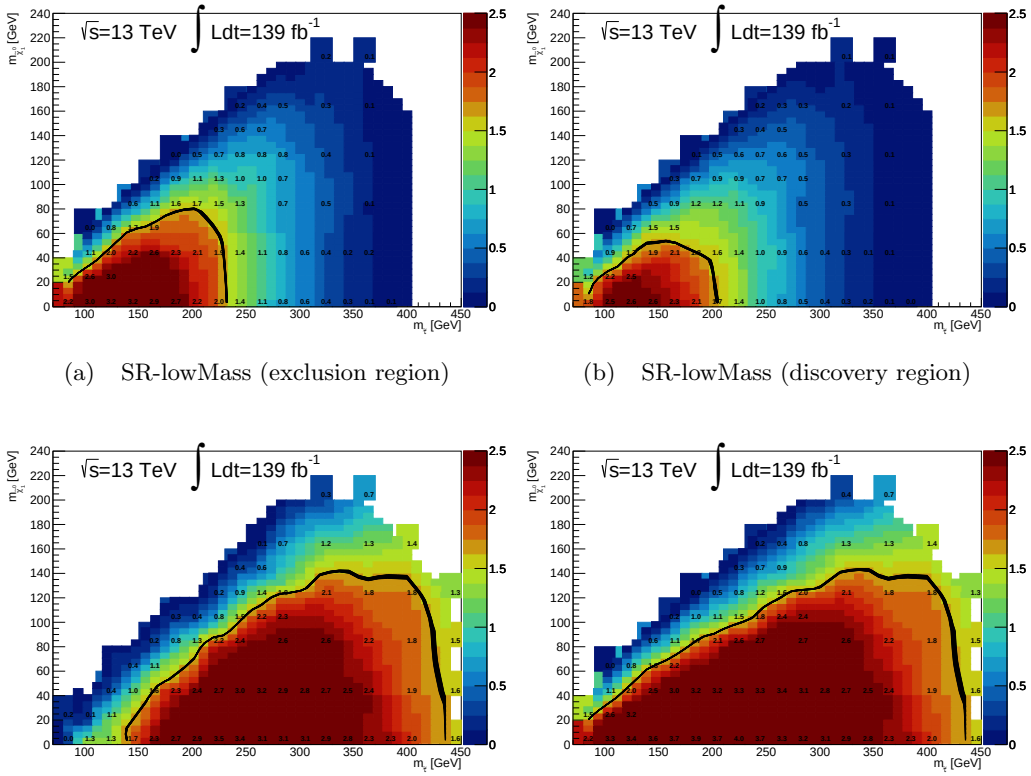
Figure 7.3.: BDT output scores. These two plots correspond to the evaluation of the BDTs on the full available dataset for the Standard Model backgrounds and a few signal benchmark models. Some of the lower bins in the plot also show data events which agree very well with the Standard Model expectation. The lower panel of these plots shows the significance in dependence of the output score.

SR highMass	SR lowMass (discovery)
2 medium τ, OS	
asymmetric ditau trigger	
3^{rd} baseline τ veto	
light lepton veto	
b-veto	
$m_{\text{T2}}(\tau_1, \tau_2) > 30\,\text{GeV}$	
$p_{\text{T}}(\tau_1) > 95\,\text{GeV}$	
$p_{\text{T}}(\tau_2) > 65\,\text{GeV}$	
$m_{\text{vis}}(\tau_1, \tau_2) > 80\,\text{GeV}$	
–	$\Delta R(\tau_1, \tau_2) < 4$
$E_{\text{T}}^{\text{miss}} > 150\,\text{GeV}$	$E_{\text{T}}^{\text{miss}} \leq 150\,\text{GeV}$
hmBDT > 0.9	lmBDT > 0.75

Table 7.4.: Signal region definitions for the discovery fit. The SR-highMass definition is the same as for the exclusion fit. "hmBDT" and "lmBDT" are short for HighMassBDT and LowMassBDT, respectively.

	Bin 1	Bin 2	Bin 3	Bin 4
Z+Jets	17.0 ± 8.4	5.8 ± 4.1	2.3 ± 1.5	$1.4^{+2.1}_{-1.4}$
W+Jets	12.7 ± 8.7	$1.0^{+7.2}_{-1.00}$	4.6 ± 4.3	6.4 ± 5.4
Top	1.99 ± 0.57	3.84 ± 0.87	5.13 ± 0.92	1.79 ± 0.57
Multiboson	3.38 ± 0.68	3.50 ± 0.66	4.00 ± 0.73	2.80 ± 0.59
Multijet	4.2 ± 2.3	9.5 ± 2.3	8.7 ± 3.2	3.1 ± 1.1
Higgs	0.0156 ± 0.0058	0.45 ± 0.43	0.0032 ± 0.0027	0.030 ± 0.026
total SM	39 ± 12	24.0 ± 8.7	24.7 ± 5.7	15.5 ± 6.0
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$	8.69 ± 0.60	11.44 ± 0.68	23.27 ± 0.98	38.32 ± 0.58
Z (+30% syst)	0.17	0.53	1.56	2.18
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (80, 1)\,\text{GeV}$	14.0 ± 1.3	13.7 ± 1.3	21.0 ± 1.6	19.0 ± 1.5
Z (+30% syst)	0.43	0.68	1.41	1.44
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)\,\text{GeV}$	0.188 ± 0.051	0.496 ± 0.086	0.88 ± 0.11	4.1 ± 1.5
Z (+30% syst)	0.00	0.00	0.00	0.14

Table 7.5.: Expected pre-fit event yields for Direct Stau SR-lowMass (exclusion). The uncertainties correspond to the statistical uncertainty only. For the calculation of the significance for the shown benchmark signal points a flat systematic uncertainty of 30% on the background yield is assumed.



(a) SR-lowMass (exclusion region)

(b) SR-lowMass (discovery region)

(c) SR-highMass

(d) Combined Exclusion SRs

Figure 7.4.: Expected exclusion range for the signal regions. For the lowMass exclusion signal region the sum of the squared significances of each bin was used to estimate the total significance. The expected significance is colour-coded and the critical value for exclusion (1.64) is marked by a black contour.

	SR-highMass	SR-lowMass (discovery)
Z+jets	0.51 ± 0.32	3.7 ± 2.5
W+jets	0.43 ± 0.36	10.9 ± 6.9
Top	0.21 ± 0.21	6.9 ± 1.1
Multiboson	1.21 ± 0.31	6.80 ± 0.94
Multijet	0.01 ± 0.21	11.8 ± 3.4
Higgs	–	0.033 ± 0.026
total SM	2.36 ± 0.66	40.2 ± 8.2
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0})$ [GeV]	$(200, 1)$: 8.90 ± 0.32	$(120, 1)$: 61.6 ± 1.6
Z (+30% syst)	2.96	2.63
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0})$ [GeV]	$(360, 1)$: 6.70 ± 0.30	$(360, 1)$: 5.01 ± 0.27
Z (+30% syst)	2.34	0.07
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0})$ [GeV]	$(440, 1)$: 4.46 ± 0.16	$(80, 1)$: 40.0 ± 2.2
Z (+30% syst)	1.62	1.79

Table 7.6.: Expected pre-fit event yields for Direct Stau SR-highMass and the SR-lowMass discovery region. The uncertainties correspond to the statistical uncertainty only. For the calculation of the significance for the shown benchmark signal points a flat systematic uncertainty of 30% on the background yield is assumed.

7.4. Background Estimation

Apart from the multijet background, which is estimated in a data-driven way, all background estimates rely on Monte Carlo simulations. The estimates of the most important backgrounds have to be evaluated in regions different to the signal regions but nevertheless in a kinematic region close to the SRs to infer the quality of the modelling of these backgrounds in the SRs (validation regions). If there are mismodelling that can be corrected by scaling the affected background this can be done in dedicated control regions. As the signal region definition relies almost exclusively on the output scores of the HighMassBDT and the LowMassBDT, these two scores are also used to identify validation regions for the respective backgrounds and control regions if necessary.

If a control region is needed, the general strategy is as follows: The control region is placed at low BDT output scores, while the corresponding VRs reside at higher LowMassBDT or HighMassBDT scores. The validation regions do not necessarily have to be statistically independent of each other. The lowMass validation regions are placed at a missing transverse energy $< 150\,\text{GeV}$ to be closer to the lowMass SR in its definition. However, due to lack of statistics at $E_{\text{T}}^{\text{miss}} > 150\,\text{GeV}$, this cut is loosened to a lower value for the highMass validation regions. Both the VRs and the CRs are separated from the signal regions by the inversion of one of the SR definition cuts. This strategy is used for the W+jets and the Top backgrounds and is illustrated in Fig. 7.5.

For the Z+jets background, no large mismodelling is found. Therefore, two Z+jets enriched validation regions are defined, which are not used in the fit. Instead, the cross section uncertainty for Z+jets is used as a nuisance parameter in the fit.

The multiboson background, besides yielding the same final state as the signal, shows also very similar kinematics to the stau signal. In addition, the production cross section of two or more weak vector-bosons is very low compared to other Standard Model processes with similar final states as Z+jets and W+jets. This makes it very difficult to isolate the multiboson background with enough purity and sufficient statistics while still keeping the signal contamination low. Nevertheless, a validation region has been designed for this background, which shows good agreement of observed and expected events. As this is the case, a control region for the normalisation of this background is not necessary for the fit and the uncertainty on the production cross section is used instead as a nuisance parameter for this background.

Background processes containing Higgs bosons pose a very small contribution in all control and signal regions. Therefore, this background is modelled by the bare Monte Carlo simulation, without defining a dedicated control or validation region. The production cross section is however taken into account as a nuisance parameter in the fit.

In the following sections, details on the definition and modelling performance in the control and validation regions for the various background categories are given.

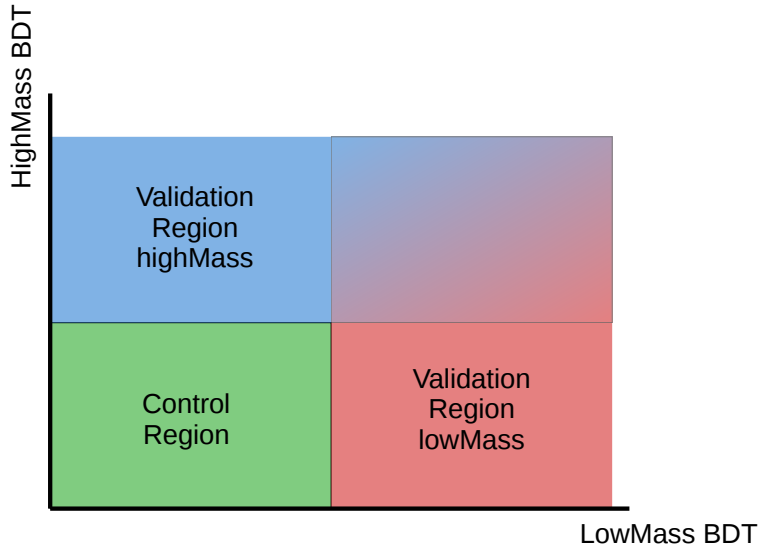


Figure 7.5.: Illustration of the CR and VR strategy. A control region is defined at low scores of both the LowMassBDT and the HighMassBDT. The highMass (lowMass) validation region is located at higher output scores of the HighMassBDT (LowMassBDT). The VRs do not necessarily have to be orthogonal to each other.

7.4.1. W+jets Estimation

The W+jets background is normalised in a control region W-CR, which is defined based on an event selection with one isolated muon and one medium τ-lepton. With this requirement, the region is orthogonal to the signal regions (exactly two τ-leptons). The control region covers the entire range of $E_{\mathrm{T}}^{\mathrm{miss}}$ while an upper cut on the LowMassBDT is applied as well as on the HighMassBDT at an output score of 0.5. The lowMass validation region (W-VR lowMass) is defined by a lower cut at LowMassBDT score at 0.5 and a maximum $E_{\mathrm{T}}^{\mathrm{miss}}$ of 150 GeV. The highMass validation region (W-VR highMass) is located at $E_{\mathrm{T}}^{\mathrm{miss}}$ values above 50 GeV and a HighMassBDT score of at least 0.5. The exact definitions are given in Tab. 7.7. Like in the first iteration of the analysis, a top-tag veto is employed in addition to the b-tag veto to suppress the Top background in these regions more effectively. The requirement on the transverse mass of the muon further improves the purity of the W+jets background. Because of the selection of an isolated muon, the improved jet rejection of the new τ-identification working point, as well as the lower cut on $m_{\mathrm{T2}}(\tau, \mu)$, the contribution of multijet events in these regions is negligible. The respective event yields for each region are listed in Tab. 7.8. Some kinematic distributions for W-CR, W-VR lowMass and W-VR highMass are shown in Fig. E.1, Fig. E.2 and Fig. E.3 in the appendix, respectively.

W-CR	W-VR lowMass	W-VR highMass
	1 medium τ, 1 signal μ, OS	
	single μ trigger	
	veto 2$^{\text{nd}}$ baseline μ and τ	
	baseline electron veto	
	b-tag veto	
	top tag veto	
	$50 < m_{\text{T}}(\mu) < 150\,\text{GeV}$	
	$m_{\text{T2}}(\tau, \mu) > 30\,\text{GeV}$	
$p_{\text{T}}(\tau) > 50\,\text{GeV}$	$\max\left[p_{\text{T}}(\tau), p_{\text{T}}(\mu)\right] > 95\,\text{GeV}$	$p_{\text{T}}(\tau) > 50\,\text{GeV}$
$p_{\text{T}}(\mu) > 50\,\text{GeV}$	$\min\left[p_{\text{T}}(\tau), p_{\text{T}}(\mu)\right] > 65\,\text{GeV}$	$p_{\text{T}}(\mu) > 50\,\text{GeV}$
$m_{\text{vis}}(\tau_1, \tau_2) > 80\,\text{GeV}$	$m_{\text{vis}}(\tau, \mu) > 120\,\text{GeV}$	–
–	$E_{\text{T}}^{\text{miss}} \leq 150\,\text{GeV}$	$E_{\text{T}}^{\text{miss}} > 50\,\text{GeV}$
LowMassBDT Score < 0.5 HighMassBDT Score < 0.5	LowMassBDT Score > 0.5	HighMassBDT Score > 0.5

Table 7.7.: W+jets CR and VR Definition. The selection is based on the requirement of one isolated muon and one medium τ-lepton, which makes the regions statistically independent of the signal regions. To suppress events from the Top background, a top tag veto is applied in addition to the b-tag veto. The top-tagging is the same as used for the first analysis iteration and is based on the m_{CT} variable as described in [112]. In this selection, the muon is most likely to originate from the W-boson while the τ-lepton is a misidentified jet. Therefore, the requirement on the transverse mass of the muon further enhances the purity of the W+jets background in these regions.

7.4.2. Top Estimation

Events containing one or more top quarks pose an important background in some of the lowMass SR bins. This background is normalised in a control region (Top-CR) which is located at LowMassBDT values between 0.25 and 0.5 and HighMassBDT scores below 0.05. For the validation regions a cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ is applied at $\leq 150\,\mathrm{GeV}$ for the lowMass validation region (Top-VR lowMass) and at $> 80\,\mathrm{GeV}$ for the highMass validation region (Top-VR highMass). Furthermore, there are lower cuts on the BDT scores for the validation regions, namely at LowMassBDT score > 0.5 for Top-VR lowMass and at HighMassBDT score > 0.05 for Top-VR highMass. All of the Top enriched regions are statistically independent of the signal regions by the inversion of the b-tag veto ($N(b\text{-tags}) \geq 1$), which also ensures the purity of the Top background in these regions. The detailed list of cuts for the top control and validation regions is given in Tab.7.9, with the respective number of events in Tab. 7.10. Some kinematic distributions for Top-CR, Top-VR lowMass and Top-VR highMass are shown in Fig. E.4, Fig. E.5 and Fig. E.6, respectively, in the appendix.

7.4.3. Z+jets Estimation

The production of a Z-boson in association with jets poses an irreducible background to the direct stau final state with two hadronically decaying τ-leptons. Two validation regions, enriched with Z+jets events, are defined. The cut on the invariant mass of the two τ-leptons in the signal regions is inverted to increase the number of Z+jets events. Additionally, an upper cut on the LowMassBDT at 0.2 (Z-VR lowMass) and on the HighMassBDT score at 0.5 (Z-VR highMass) are placed together with a cut on the missing transverse energy. A detailed definition of the Z+jets validation regions is given in Tab. 7.11. The event yields in the two regions are listed in Tab. 7.12. Some kinematic

	W-CR	W-VR lowMass	W-VR highMass
Z+jets	249 ± 50	5.4 ± 2.1	$0.06^{+0.16}_{-0.06}$
W+jets	4330 ± 430	337 ± 49	16.1 ± 4.9
Top	105.0 ± 4.0	20.0 ± 1.8	1.59 ± 0.67
Multiboson	145.1 ± 5.0	22.8 ± 1.7	1.61 ± 0.49
Higgs	3.8 ± 1.4	0.77 ± 0.77	–
total SM	4830 ± 430	386 ± 49	19.35 ± 5.01
Data	4138	357	16
Data/MC ratio	0.856 ± 0.078	0.93 ± 0.13	0.83 ± 0.30
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)\,\mathrm{GeV}$	0.602 ± 0.082	1.52 ± 0.13	0.757 ± 0.092
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\mathrm{GeV}$	2.64 ± 0.33	2.16 ± 0.30	0.53 ± 0.15
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)\,\mathrm{GeV}$	–	0.149 ± 0.047	0.131 ± 0.042

Table 7.8.: Pre-fit event yields in W+jets CR and VRs. The ratio of data and simulated events in this region is around 0.9. The normalisation of W+jets is performed in the W-CR. The given uncertainties correspond to the statistical uncertainties only. The multijet background is removed from this table as its contribution is negligible in these regions.

Top-CR	Top-VR lowMass	Top-VR highMass
2 medium τ, OS		
asymmetric ditau trigger		
3^{rd} baseline τ veto		
light lepton veto		
≥ 1 b-tag		
$m_{\text{T2}}(\tau_1, \tau_2) > 30\,\text{GeV}$		
$p_{\text{T}}(\tau_1) > 95\,\text{GeV}$		
$p_{\text{T}}(\tau_2) > 65\,\text{GeV}$		
$m_{\text{vis}}(\tau_1, \tau_2) > 80\,\text{GeV}$		–
–	$E_{\text{T}}^{\text{miss}} \leq 150\,\text{GeV}$	$E_{\text{T}}^{\text{miss}} > 80\,\text{GeV}$
$0.25 <$ LowMassBDT Score < 0.5 HighMassBDT Score < 0.05	LowMassBDT Score > 0.5	HighMassBDT Score > 0.05

Table 7.9.: Top CR and VR Definition. The regions are orthogonal to the signal regions by inversion of the b-tag veto. The invariant mass cut primarily reduces contributions from Z+jet.

	T-CR	T-VR lowMass	T-VR highMass
Z+jets	1.60 ± 1.2	0.20 ± 0.53	0.76 ± 0.23
W+jets	$0.8^{+1.0}_{-0.8}$	2.7 ± 1.4	1.7 ± 1.3
Top	36.2 ± 2.5	29.3 ± 2.2	17.2 ± 1.7
Multiboson	0.230 ± 0.080	0.15 ± 0.15	0.36 ± 0.10
Multijet	$0.16^{+0.86}_{-0.16}$	$0.00^{+0.66}_{-0.00}$	$0.23^{+0.63}_{-0.23}$
Higgs	0.0578 ± 0.0097	0.0412 ± 0.0070	0.170 ± 0.032
total SM	39.1 ± 2.8	32.5 ± 2.7	20.4 ± 2.2
Data	34	25	17
Data/MC ratio	0.87 ± 0.16	0.77 ± 0.17	0.83 ± 0.22
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)\,\text{GeV}$	0.096 ± 0.031	0.506 ± 0.076	0.645 ± 0.085
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$	0.30 ± 0.11	0.89 ± 0.20	0.46 ± 0.14
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)\,\text{GeV}$	0.031 ± 0.024	0.068 ± 0.031	0.314 ± 0.068

Table 7.10.: Pre-fit event yields in Top CR and VRs. The ratio of observed to expected Standard Model events is around 0.85. The T-CR is used to normalise the Top background in the fit. The given uncertainties correspond to the statistical uncertainties only.

distributions of Z-VR lowMass are shown in Fig. E.7 and in Fig. E.8 for Z-VR highMass.

Z-VR lowMass	Z-VR highMass
2 medium τ, OS	
asymmetric ditau trigger	
3^{rd} baseline τ veto	
light lepton veto	
b-veto	
$m_{\text{T2}}(\tau_1, \tau_2) > 30\,\text{GeV}$	
$p_{\text{T}}(\tau_1) > 95\,\text{GeV}$	
$p_{\text{T}}(\tau_2) > 65\,\text{GeV}$	
$m_{\text{vis}}(\tau_1, \tau_2) < 100\,\text{GeV}$	$m_{\text{vis}}(\tau_1, \tau_2) < 120\,\text{GeV}$
$E_{\text{T}}^{\text{miss}} \leq 150\,\text{GeV}$	$E_{\text{T}}^{\text{miss}} > 100\,\text{GeV}$
LowMassBDT Score < 0.2	HighMassBDT Score < 0.5

Table 7.11.: Z+jets VR Definitions. A high purity in Z+jets events is achieved by setting an upper cut on the invariant mass of the two τ-leptons. The lowMass Z-VR is defined at low LowMassBDT scores and the highMass Z-VR at low HighMassBDT scores.

7.4.4. Multiboson Estimation

Events with more than one vector boson are the dominant background in the highMass signal region. A validation region MB-VR is defined based on a selection with two tight τ-leptons. The MB-VR lies at HighMassBDT scores below 0.9 and has an inverted cut on the invariant mass of the two τ-leptons compared to the signal regions. Therefore, the signal contamination in this region is very small. Because of the requirement of the two tight τ-leptons, events containing one or more fake τ-leptons are largely rejected. This leads to the W+jets and multijet backgrounds being negligible in this region. The definition of MB-VR is given in Tab. 7.13 and the event yields are listed in Tab. 7.14. Some kinematic distributions in this region are given in Fig. E.9 in the appendix.

7.4.5. Multijet Estimation

The estimation of the multijet background is performed in two slightly different ways, which also serve two slightly different purposes. One of them is via the ABCD method, as it has been already described in Sec. 6.3.1. In this version of the ABCD method, the estimation is done based on $m_{\text{T2}}(\tau_1, \tau_2)$ and τ-identification working point and tau charge requirements. A loose τ-identification working point (excluding medium and tight working points) for oppositely charged τ-leptons in a logical "or" with the requirement of the two τ-leptons to have the same charge sign (passing a τ-identification working point of at least "loose") is defining the multijet enriched control and validation regions, CR-B, VR-E and VR-A. For the regions CR-C, VR-F and preSR-D two at least medium τ-leptons are required with oppositely signed charge. The control regions in which the transfer factor is calculated are further defined by $10 < m_{\text{T2}}(\tau_1, \tau_2) < 20\,\text{GeV}$. CR-A and

	Z-VR lowMass	Z-VR highMass
Z+jets	548 ± 15	28.2 ± 4.5
W+jets	1.82 ± 0.80	0.49 ± 0.30
Top	0.66 ± 0.39	0.60 ± 0.29
Multiboson	17.55 ± 0.87	7.81 ± 0.74
Multijet	14.2 ± 2.9	0.00 ± 0.49
Higgs	5.24 ± 0.95	2.32 ± 0.81
total SM	587 ± 15	39.5 ± 4.6
Data	565	34
Data/MC ratio	0.962 ± 0.047	0.86 ± 0.18
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)$ GeV	0.099 ± 0.033	0.407 ± 0.067
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)$ GeV	0.164 ± 0.082	0.43 ± 0.14
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)$ GeV	0.034 ± 0.021	0.262 ± 0.060

Table 7.12.: Pre-fit event yields in Z+jets VRs. The Monte Carlo simulation yields a sufficiently good modelling of the Z+jets background, as can be seen from the kinematic distributions in Sec. E.3 in the appendix. The quoted uncertainties are statistical uncertainties only.

Multiboson-VR
2 tight τ-leptons (OS)
3rd baseline tau veto
b-veto
asymmetric ditau trigger
$m_{\mathrm{T2}}(\tau_1, \tau_2) > 10\,\mathrm{GeV}$
$m_{\mathrm{vis}}(\tau_1, \tau_2) < 120\,\mathrm{GeV}$
$m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_1) > 270\,\mathrm{GeV}$
$\Delta R(\tau_1, \tau_2) < 4$
$0.02 <$ HighMassBDT Score < 0.9

Table 7.13.: Multiboson VR Definition. To reduce background events with jets misidentified as τ-leptons, the required τ-identification working point is raised to "tight". It is furthermore located at intermediate and lower values of the HighMassBDT score and therefore orthogonal to the highMass signal region. This is not strictly necessary for a validation region in general, but this way the signal contamination in this region is also reduced. Furthermore, the invariant mass cut is inverted compared to the signal regions.

	Multiboson-VR
Z+jets	6.8 ± 1.2
W+jets	–
Top	0.261 ± 0.053
Multiboson	7.61 ± 0.72
Multijet	–
Higgs	1.31 ± 0.65
total SM	16.0 ± 1.5
Data	16
Data/MC ratio	1.00 ± 0.27
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)$ GeV	0.261 ± 0.045
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)$ GeV	0.045 ± 0.045
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)$ GeV	0.380 ± 0.072

Table 7.14.: Pre-fit event yields in the Multiboson VR. The requirement of the two τ-leptons passing the tight identification working point reduces the W+jets and multijet background to a negligible level. The contribution of Z+jets is almost as large as the multiboson contribution. It is however very difficult to separate these backgrounds further without losing too many events in this region. As the Z+jets simulation itself showed good performance in the Z-VRs, the very good agreement of observed and simulated events also testifies for a good multiboson modelling.

the preliminary SR-D are located at values of $m_{\mathrm{T2}}(\tau_1, \tau_2) > 30\,\mathrm{GeV}$. In-between are the validation regions VR-E and VR-F: $20 < m_{\mathrm{T2}}(\tau_1, \tau_2) < 30\,\mathrm{GeV}$. The detailed list of cuts for the various ABCD-regions is given in Tab. 7.15. In each region, except pre-SR-D, the multijet contribution is assumed to be the difference of the observed data to the Monte Carlo simulated Standard Model background. The transfer factor is defined as the ratio of the multijet contribution in CR-C and CR-B. If $m_{\mathrm{T2}}(\tau_1, \tau_2)$ and the τ-identification and sign requirements were completely uncorrelated, the exact same transfer factor would apply to CR-A and pre-SR-D, and to VR-E and VR-F as well. The correlation is shown in Fig. 7.7(a) by drawing the normalised $m_{\mathrm{T2}}(\tau_1, \tau_2)$ distributions for the multijet estimate (i.e. the difference in yields from data and Standard Model simulation) in the loose τ-identification or same sign regions (CR-B, VR-E and CR-A) and in the medium τ-identification and opposite sign regions (CR-C, VR-F, and small parts of pre-SR-D). If the ratio of these (normalised) distributions is compatible with unity, the correlation is small and it is sufficient to include it as a systematic uncertainty in the fit. The systematic uncertainty on the transfer factor due to correlation is estimated by calculating the transfer factor again in VR-E and VR-F and evaluating the difference. In addition, a statistical uncertainty on the transfer factor is also taken into account. The resulting transfer factors are given in Tab. 7.16. The modelling of the resulting distributions after deriving the multijet estimation with the ABCD method can be checked not only in the validation region, but also in pre-SR-D, as this preliminary signal region is defined in a rather loose way such that the signal contamination is very small. The corresponding plots are shown in the appendix in Sec. E.5.2 for pre-SR-D and in Sec. E.5.3 for VR-F. The multijet estimate used for the distributions in VR-F has been derived using the transfer factor calculated from CR-B and CR-C as well. A detailed table with all the event yields per Standard Model process per region is given in Sec. E.5.1.

Although the modelling appears good in the plots, this method has one caveat: By estimating the multijet contribution in each region as the difference of data and Standard Model simulation, all Monte Carlo events in CR-A enter the multijet estimate for pre-SR-D with a negative sign. The multijet background should be part of the background class in the training of the BDTs. As BDTs change the weights of the input events during the training, an event weight that is already negative from the very beginning of the training can degrade the BDTs performance notably.[4] To avoid this problem, the ABCD method is modified into a "simplified" version: Instead of applying the transfer factor directly to the difference of data and simulation events, the data events are scaled to represent the number of multijet events. The transfer factor is then applied to the scaled data events. The weight factor used to scale the data in CR-A is calculated as follows:

$$W = \frac{\mathrm{Data(A)} - \mathrm{MC(A)}}{\mathrm{Data(A)}}. \tag{7.2}$$

A condition for this to work is that the $m_{\mathrm{T2}}(\tau_1, \tau_2)$ distributions for data and the data to simulation difference deviate not too much from each other, i.e. that they show the same shape. This is checked in the normalised plot in Fig. 7.7(b). The ratio of the two normalised distributions is compatible with one, and therefore it is legitimate to use a single weight factor for the scaling of the data in CR-A. As a consequence of using scaled data events, the multijet estimate does not introduce any negatively weighted events to the BDT training. The good modelling performance of this slightly modified ABCD

[4] A similar argument holds for events with very large weights. Events with weights > 50 are therefore excluded from the training but re-included in the evaluation.

method can be verified by the kinematic distributions in pre-SR-D and VR-F shown in the appendix in Sec. E.5.4. The simplified ABCD method is illustrated and summarized in Fig. 7.6. For the final analysis results, however, the original version of the ABCD method is used to avoid introducing additional systematic uncertainties.

When evaluating the background contributions to the four bins of the SR-lowMass it can be seen that multijet events pose one of the most important backgrounds in all of these bins. Therefore, the modelling should be more thoroughly tested in a multijet enriched region. As VR-F does not yield a good purity in multijet events and is also already used to derive the systematic uncertainty on the multijet estimate, an additional validation region for multijet events is defined in Tab. 7.17. The respective number of events for background and observed data as well as for a few signal benchmark points are given in Tab. 7.18. From this table as well as from the corresponding kinematic distributions in the appendix in Sec. E.5.6 good agreement of the expected background and data can be observed.

<table>
<tr><td>CR-A</td><td>VR-E</td><td>CR-B</td><td>pre-SR-D</td><td>VR-F</td><td>CR-C</td></tr>
<tr><td colspan="6">asymmetric ditau trigger without plateau cuts
2 τ-leptons passing at least loose τ-ID
light lepton veto</td></tr>
<tr><td colspan="3">2 loose (not medium) τ-leptons (OS) OR 2 loose τ (SS)</td><td colspan="3">2 τ-leptons passing at least medium τ-ID</td></tr>
<tr><td>$m_{\mathrm{T2}} > 30\,\mathrm{GeV}$</td><td>$20 < m_{\mathrm{T2}} < 30\,\mathrm{GeV}$</td><td>$10 < m_{\mathrm{T2}} < 20\,\mathrm{GeV}$</td><td>$m_{\mathrm{T2}} > 30\,\mathrm{GeV}$</td><td>$20 < m_{\mathrm{T2}} < 30\,\mathrm{GeV}$</td><td>$10 < m_{\mathrm{T2}} < 20\,\mathrm{GeV}$</td></tr>
</table>

Table 7.15.: Region definitions for the ABCD method. $m_{\mathrm{T2}}(\tau_1, \tau_2)$, the τ-identification working points and the relative sign of the τ-charges are used to define the regions.

Transfer factor (calculated from CR-B and CR-C)	0.158 ± 0.021
Transfer factor (calculated from VR-E and VR-F)	0.123 ± 0.025
Weight factor (calculated from CR-A)	0.773 ± 0.014

Table 7.16.: Transfer factors from CRs and VRs in the ABCD method and the weight factor for the simplified ABCD method with their respective statistical uncertainty. The transfer factor calculated in VR-E and VR-F is used to estimate the systematic uncertainty on the transfer factor due to correlation of the requirements in the ABCD region definition.

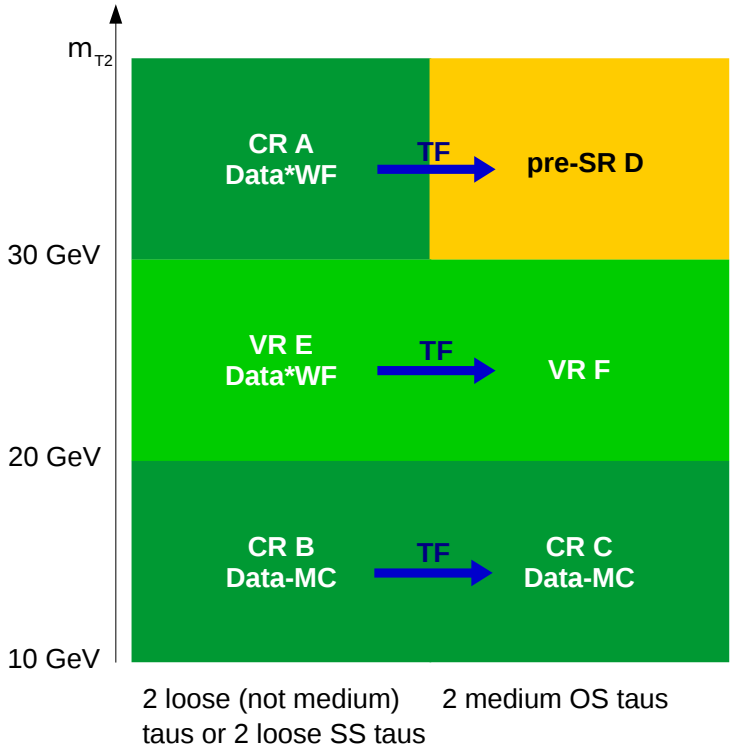
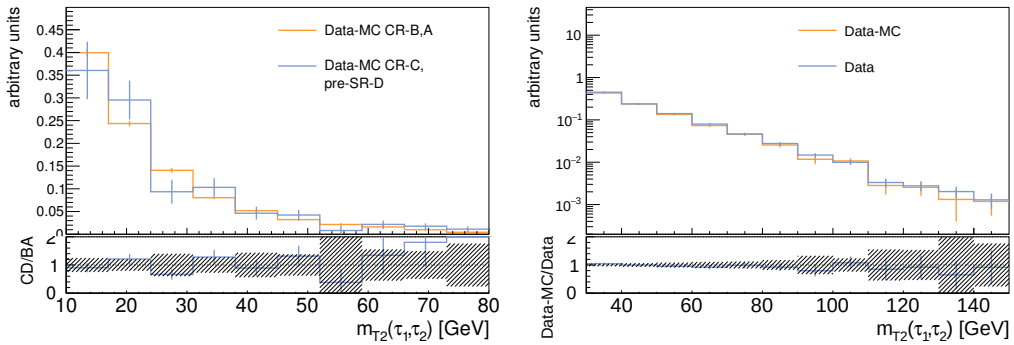


Figure 7.6.: Illustration of the simplified ABCD Method. Instead of applying the transfer factor to the difference of data and simulation events, the data is scaled by a weight factor to the expected number of multijet events in CR-A. This way, the performance of the BDT training is improved by avoiding a large number of events with negative weights. In the regions CR-B and CR-C the difference of data and simulated events is still used for the calculation of the transfer factor. A separate weight factor is derived for VR-E in the same way as for CR-A before determining the multijet estimate in VR-F by applying the transfer factor.



(a) Correlation check (b) Shape check

Figure 7.7.: Correlation and shape checks for the (simplified) ABCD method. The left plot shows the correlation of the requirements on m_{T2} and charge sign and τ-identification used to define the regions for the ABCD method. It is found to be negligible such that it enters the analysis as a systematic uncertainty. The right plot shows the agreement of the $m_{\mathrm{T2}}(\tau_1, \tau_2)$ shapes for data and the difference of data and simulated events in CR-A, thus justifying the use of a single weight factor to scale data in the simplified ABCD method.

Multijet-VR
2 medium τ (OS)
asymmetric ditau trigger
light lepton veto
b-veto
$m_{\mathrm{vis}}(\tau_1, \tau_2) > 120\,\mathrm{GeV}$
$m_{\mathrm{T2}}(\tau_1, \tau_2) > 30\,\mathrm{GeV}$
$\Delta R(\tau_1, \tau_2) > 3$
$E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$
LowMass BDT score < 0.4

Table 7.17.: Multijet-VR definition. In addition to the regions used in the ABCD method, a validation region enriched with multijet events is defined. It serves to check the modelling of the multijet estimate as this background is dominant in the SR-lowMass bins.

	Multijet-VR
Z+jets	8.5 ± 6.4
W+jets	20.0 ± 8.1
Top	1.92 ± 0.55
Multiboson	1.75 ± 0.41
Multijet	58.6 ± 3.6
Higgs	0.0082 ± 0.0050
total SM	91 ± 11
Data	84
Data/MC ratio	0.92 ± 0.15
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)$ GeV	0.397 ± 0.068
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)$ GeV	1.10 ± 0.21
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)$ GeV	0.035 ± 0.021

Table 7.18.: Expected pre-fit event yields for the multijet-VR. There is good agreement of observed and expected events with a ratio of data and expected events compatible with unity. Kinematic distributions for this region can be found in the appendix in Sec. E.5.6.

	Normalisation factor
μ_{Top}	0.867 ± 0.213
μ_{W}	0.844 ± 0.089

Table 7.19.: Normalisation factors for W+jets and Top as derived during the background-only fit.

7.5. Fit Setup and Results

7.5.1. Systematic Uncertainties

In principle, the same sources of systematic uncertainties need to be taken into account as in the first direct stau search (cf. Sec. 6.4). Apart from an updated scheme for the jet energy resolution systematics, the same experimental systematic uncertainties are considered. The theoretical uncertainties are included as flat uncertainties in the fit and will be replaced in the further progress of the analysis. The magnitudes of these flat uncertainties are motivated by the theoretical uncertainty evaluation for the first analysis iteration.

7.5.2. Background-Only Fit

Only the W-CR and T-CR defined in Sec. 7.4.1 and Sec. 7.4.2 are used for the fit, while the signal regions are treated as validation regions. The normalisation factors for the W+jets and the Top backgrounds obtained from the fit are given in Tab. 7.19. Additional information on the fit result can be found in the appendix in Sec. F.1. The results of the fit parameters are applied in all validation regions and signal regions. The agreement of the observed and the expected number of events after the fit is checked for every region. All the expected event yields in the validation regions agree with the observed ones within $\pm 0.5\sigma$ or even less as can be seen in Fig. 7.8. The signal regions, though treated as validation regions in the background-only fit, are still blinded. The exact event yields can also be seen in the appendix in Tab. F.1 for the Top CR and VRs, Tab. F.2 for the W+jets CR and VRs, Tab. 7.20 and Tab. 7.21 for the signal regions and Tab. F.3 for the remaining validation regions. The most important systematic uncertainties after the fit for each region are shown in Tab. 7.22.

7.5.3. Model Independent Cross Section Limits

To set model independent limits, an arbitrary signal parameter μ_{sig} is included. An upper limit scan is performed in a certain range of μ_{sig} and the CL_s is evaluated for each scan point. This way, an upper limit for any additional (non-Standard Model) process can be derived. The values of μ_{sig} at which $CL_s > 0.05$ corresponds to the number of additional events in the signal region that can be explained by a statistical fluctuation of the Standard Model background. This number of events can then be converted to an effective cross section σ_{vis} that can be excluded. Below $CL_s = 0.05$, the number of additional events can be excluded. As the actual data taken by ATLAS in 2015-2018 is not yet viewed in this analysis, only the background expectation from the background-only fit is used. The determination of model-independent limits is performed for SR-highMass and the

SR-lowMass (exclusion region)	Total	Bin 1	Bin 2	Bin 3	Bin 4
Post-fit background events	98 ± 27	37 ± 19	23 ± 16	23 ± 10	14.3 ± 6.8
Post-fit Z+jets events	27 ± 19	17 ± 13	$5.8^{+8.0}_{-5.8}$	$2.3^{+5.3}_{-2.3}$	$1.4^{+2.2}_{-1.4}$
Post-fit W+jets events	20.7 ± 8.9	10.7 ± 9.9	1^{+12}_{-1}	$3.8^{+6.0}_{-3.8}$	5.4 ± 3.8
Post-fit Top events	11.0 ± 3.8	1.72 ± 0.91	3.3 ± 1.6	4.4 ± 1.8	1.55 ± 0.88
Post-fit Multiboson events	13.7 ± 4.6	3.4 ± 1.8	3.5 ± 2.0	4.0 ± 1.6	2.8 ± 1.4
Post-fit Higgs events	0.50 ± 0.37	$0.016^{+0.027}_{-0.016}$	0.45 ± 0.36	$0.003^{+0.026}_{-0.003}$	$0.030^{+0.031}_{-0.030}$
Post-fit Multijet	25.4 ± 7.8	4.2 ± 1.7	9.4 ± 4.2	8.7 ± 3.0	3.1 ± 1.4
Pre-fit Standard Model events	103 ± 29	39 ± 21	24 ± 18	25 ± 11	15.5 ± 7.5
Pre-fit Z+jets events	27 ± 19	17 ± 13	$5.8^{+8.0}_{-5.8}$	$2.3^{+5.4}_{-2.3}$	$1.4^{+2.2}_{-1.4}$
Pre-fit W+jets events	25 ± 10	13 ± 12	1^{+14}_{-1}	$4.6^{+7.1}_{-4.6}$	6.4 ± 4.4
Pre-fit Top events	12.8 ± 3.1	1.99 ± 0.94	3.8 ± 1.6	5.1 ± 1.7	1.79 ± 0.92
Pre-fit Multiboson events	13.7 ± 4.6	3.4 ± 1.8	3.5 ± 2.0	4.0 ± 1.6	2.8 ± 1.4
Pre-fit Higgs events	0.50 ± 0.37	$0.016^{+0.027}_{-0.016}$	0.45 ± 0.36	$0.003^{+0.026}_{-0.003}$	$0.030^{+0.031}_{-0.030}$
Pre-fit Multijet	25.4 ± 7.9	4.2 ± 1.7	9.5 ± 4.2	8.7 ± 3.0	3.1 ± 1.4

Table 7.20.: Events yields pre- and post-fit for the SR-lowMass bins (exclusion region). The given uncertainties correspond to the combined statistical and systematic uncertainties.

	SR-lowMass (discovery region)	SR-highMass
Post-fit background events	38 ± 14	2.27 ± 0.85
Post-fit Z+jets events	$3.8^{+6.8}_{-3.8}$	0.51 ± 0.35
Post-fit W+jets events	9.2 ± 7.9	0.36 ± 0.13
Post-fit Top events	6.0 ± 2.3	0.18 ± 0.12
Post-fit Multiboson events	6.8 ± 2.5	1.21 ± 0.58
Post-fit Higgs events	0.033 ± 0.016	–
Post-fit Multijet events	11.8 ± 3.9	0.010 ± 0.0029
Pre-fit Standard Model events	40 ± 15	2.36 ± 0.88
Pre-fit Z+jets events	$3.7^{+6.9}_{-3.7}$	0.51 ± 0.35
Pre-fit W+jets events	10.9 ± 9.3	0.43 ± 0.14
Pre-fit Top events	6.9 ± 2.1	0.21 ± 0.13
Pre-fit Multiboson events	6.8 ± 2.5	1.21 ± 0.58
Pre-fit Higgs events	0.033 ± 0.016	–
Pre-fit Multijet events	11.8 ± 3.9	0.0079 ± 0.0030

Table 7.21.: Events yields pre- and post-fit for SR-highMass and the SR-lowMass discovery region. The given uncertainties correspond to the combined statistical and systematic uncertainties.

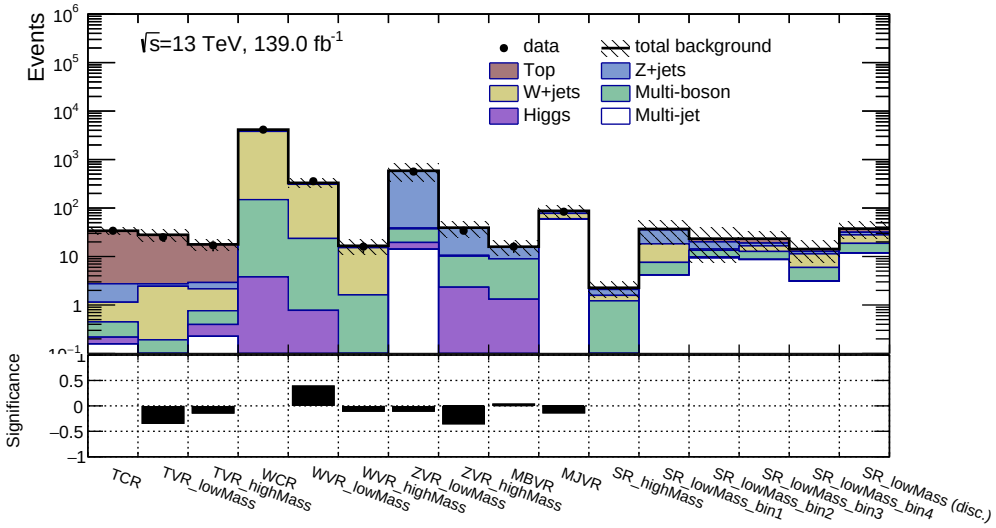


Figure 7.8.: Post-fit event yields for all SRs, CRs and VRs. The observed data has not yet been unblinded for the signal regions. The overall agreement of observed and expected events in the validation regions is within $\pm 1\sigma$ or even less. In the control regions, the observed and expected event yields agree perfectly per construction, due to the normalisation performed on the respective backgrounds in these regions.

Source of uncertainty	SR-highMass [%]	SR-lowMass (excl.) [%]	SR-lowMass (disc.) [%]
Statistical uncertainty of simulation	28	12	20
Multiboson theory uncertainty	16	4	5
Z+jets theory uncertainty	9	11	4
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft term resolution and scale	9	9	16
Tau energy scale and resolution	8	6	7
Multijet estimation	< 1	6	7
Jet energy scale and resolution	5	5	14
Tau Identification	4	2	1
Total	38	28	38

Table 7.22.: Dominant uncertainties in the signal regions after the fit. It has to be noted that due to correlations among the systematic uncertainties, the individual uncertainty sources generally might not result in the total uncertainty when added in quadrature.

discovery region SR-lowMass separately. The results are shown in Tab. 7.23. The upper limit scans are shown in the appendix in Sec. F.2.

	Expected σ_{vis}^{95} [fb]
SR-lowMass (discovery)	$0.199^{+0.059}_{-0.046}$
SR-highMass	$0.032^{+0.021}_{-0.009}$

Table 7.23.: Result of the model independent upper limit scan. The upper limit at 95% CL for the effective non-Standard Model cross section is given separately for SR-highMass and the lowMass discovery signal region.

7.6. Interpretation: Model Dependent Limits

The observed data events are not yet unblinded in this analysis. However, if there is no reason to reject the background-only hypothesis, also the signal+background hypothesis can be tested for particular stau pair production models. The signal is parameterized by the signal strength μ_{sig} in the fit. An upper limit scan is performed by varying the signal strength parameter and checking the CL_s value for each of these variations. If the CL_s value drops below 0.05 at a signal strength < 1, the considered signal model can be excluded at 95% CL. The CL_s values and the amount of excluded points can be illustrated in the plane of stau mass and the mass of the lightest neutralino, which is shown in Fig. 7.9. In addition, separate upper limit scans have been performed for the separate stau eigenstates. The exclusion contour for $\tilde{\tau}_{\text{L}}$ pair production is given in Fig. 7.10 and in Fig. 7.11 for $\tilde{\tau}_{\text{R}}$ pair production. In these figures, the dashed black line shows the expected exclusion contour, assuming that the background estimate is exactly correct. The yellow band corresponds to the $\pm 1\sigma$ uncertainty for the exclusion limit.

In comparison to the first search for stau pair production, the exclusion contour now covers stau masses as low as 80 GeV, which means that the highest masses that LEP could exclude for $\tilde{\tau}_{\text{R}}$ could come into exclusion reach with LHC run 2 data for the combined $\tilde{\tau}_{\text{L}}$ and $\tilde{\tau}_{\text{R}}$ pair production. The exclusion power is also increased for higher neutralino masses and stau masses. For comparison, the observed exclusion contour for the first analysis iteration is also shown as shaded grey area in the plot. Concerning the $\tilde{\tau}_{\text{L}}$ eigenstate, the exclusion contour has also extended notably towards lower stau masses compared to the result of the first analysis iteration, while also covering stau masses up to 350 GeV for a near-massless lightest neutralino. Because of the much smaller cross sections, it is more difficult to achieve exclusion sensitivity for the pair production of $\tilde{\tau}_{\text{R}}$ separately than for $\tilde{\tau}_{\text{L}}$. A CL_s value of < 0.05 for signal strengths below unity was, however, derived for a few models with a very light neutralino and stau masses between 120 GeV and 230 GeV, although from Fig. 7.11 no expected exclusion of these points within -1σ is visible. However, a close look at the upper limit scans in Fig. 7.12 shows that there is an expected exclusion within $\pm 1\sigma$ as the graph for $+1\sigma$ traverses the 0.05 CL_s line at a signal strength of approximately 0.9 for stau masses of 120 GeV and 160 GeV with near to massless neutralino. This is not visible of Fig. 7.11 as it is "smoothed out" in the interpolation process that derives the contours from the individual CL_s values of the considered models.

The upper limit scans for some benchmark points for the combined and separate eigenstate scenarios can be viewed in the appendix in Sec. F.3.

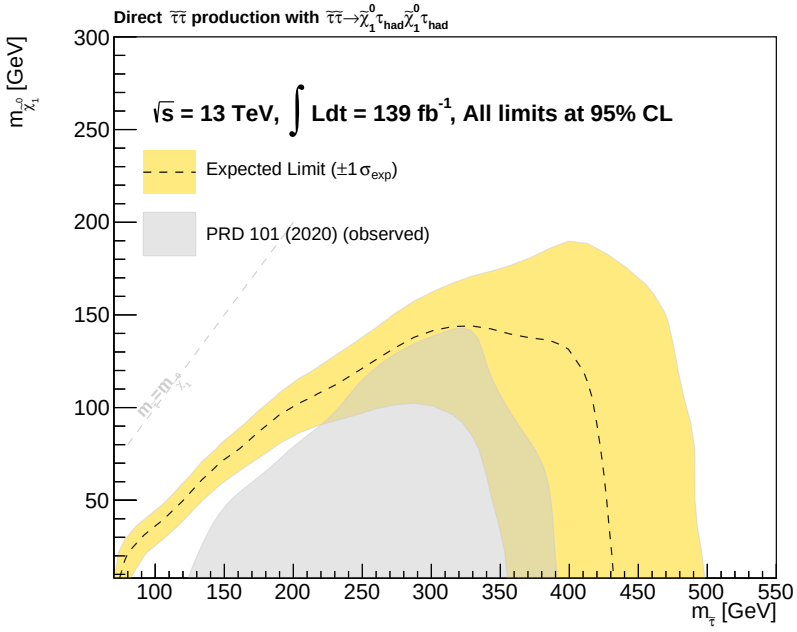


Figure 7.9.: This plot shows the expected exclusion power of the second search for stau pair production with the dashed black line. The yellow band corresponds to the $\pm1\sigma$ uncertainty. The shaded grey area corresponds to the observed exclusion contour of the first stau search as shown in Fig. 6.6 and in [51]. (Due to a small and not significant excess in data in the lowMass SR the observed exclusion in the low stau mass area was weaker than expected for the first stau search.) Comparing the two contours, the sensitivity to stau pair production could be extended not only to lower stau masses and models with a smaller mass difference between stau and lightest neutralino, but also to higher stau masses.

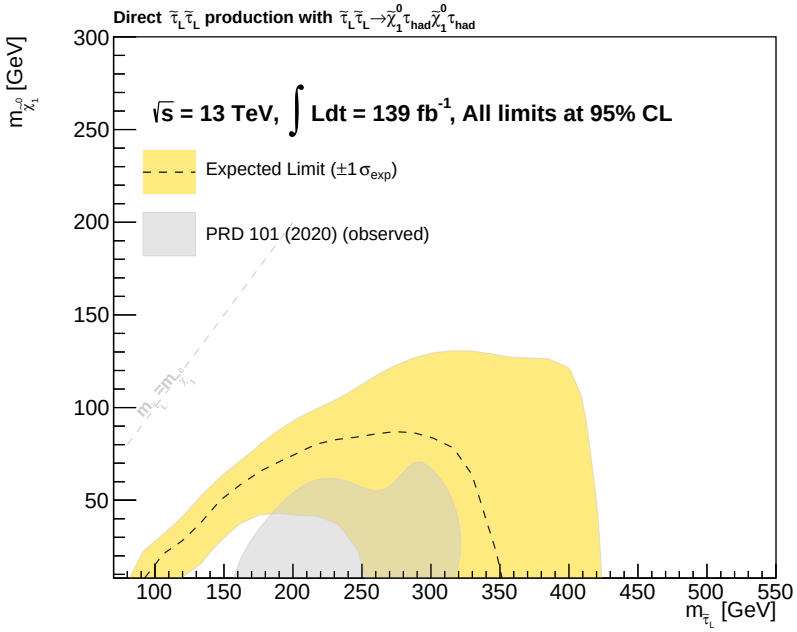


Figure 7.10.: Exclusion contour for the production of $\tilde{\tau}_L$ pairs. This plot shows the expected exclusion power of the second search for $\tilde{\tau}_L$ pair production with the dashed black line. The yellow band corresponds to the $\pm 1\sigma$ uncertainty. The shaded grey area corresponds to the observed exclusion contour of the first stau search as shown in Fig. 6.7 and in [51]. Comparing the two contours, the sensitivity to stau pair production could be extended notably to lower stau masses.

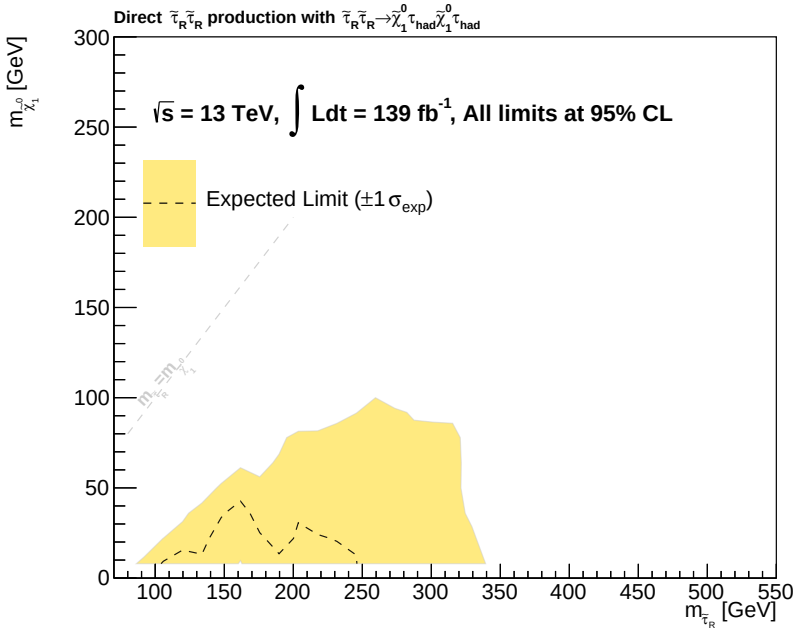
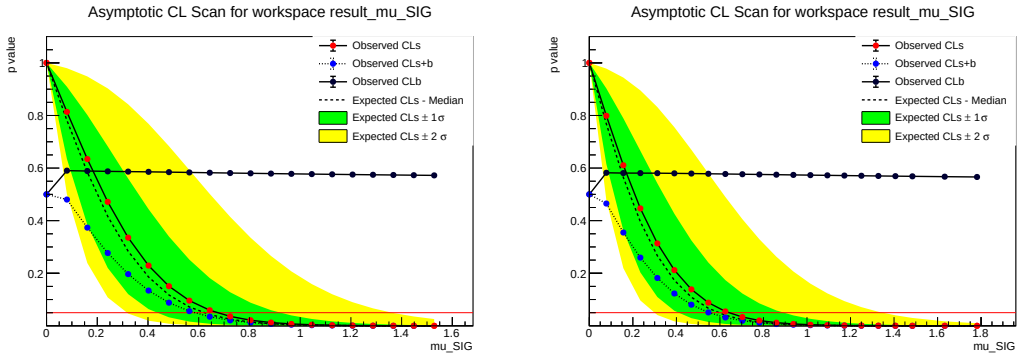


Figure 7.11.: Exclusion contour for the production of $\tilde{\tau}_\mathrm{R}$ pairs. This plot shows the expected exclusion power of the second search for $\tilde{\tau}_\mathrm{R}$ pair production with the dashed black line. The yellow band corresponds to the $\pm 1\sigma$ uncertainty. Up to now, separate exclusion sensitivity for the pair production of the $\tilde{\tau}_\mathrm{R}$ eigenstate could not be achieved. In this analysis, stau masses of 120 GeV to 230 GeV are narrowly expected to be excluded for nearly massless lightest neutralinos. The $\pm 1\sigma$ band covers the entire area under the contour. The individual upper limit scans for stau masses of 120 GeV and 160 GeV in Fig. 7.12 show, however, that there is actually an expected exclusion within $\pm 1\sigma$. The reason for this being not visible is the smoothing of the curves in the interpolation process which derives the contours from individual CL_s values of the models.



(a) $(m_{\tilde{\tau}_R}, m_{\tilde{\chi}1\,0}) = (120, 1)$ GeV (b) $(m_{\tilde{\tau}_R}, m_{\tilde{\chi}1\,0}) = (160, 1)$ GeV

Figure 7.12.: Upper limit scans for a $\tilde{\tau}_R$-mass of 120 and 160 GeV. These plots show that for these two models of $\tilde{\tau}_R$ pair production there is expected exclusion sensitivity within the $\pm 1\sigma$ uncertainty (green band).

7.7. Short Term Outlook

The next step in the presented analysis would obviously be to look at the observed ATLAS data taken in 2015-2018. The main reason for postponing this step is a bug in the merging step (cf. 4.2.2) of the simulation software affecting the signal simulation for this analysis iteration. The first iteration is not affected. This bug has only been discovered shortly before the submission of this thesis and its impact is currently under investigation. First studies on truth-level concerning this issue are shown in the appendix in Sec. G. From these studies it can be concluded that a large impact by a correction of this bug is not to be expected for models with higher stau masses. However, the sensitivity for low stau masses could potentially be decreased. To fully understand the magnitude of this effect, all of the signal simulations have to be re-generated and newly evaluated, which is beyond the time scale of this thesis.

8. Conclusion

This thesis presented two approaches to search for the direct pair production of the supersymmetric partners of the τ-lepton in events with two hadronically decaying τ-leptons, and their respective (expected) results. For both analyses the full dataset taken by the ATLAS detector at the LHC in 2015-2018 was used.

The first approach relies on the so-called cut-and-count signal region optimization method, which scans through different combinations of cut variables and values, thus maximizing the significance for the chosen benchmark signal models. This way, two signal regions have been designed – one sensitive to signal models with low stau masses (SR-lowMass) and the other with larger sensitivity for models with high stau masses (SR-highMass). In the definition of these regions, the m_{T2} variable played a crucial role. The W+jets background is normalized in a control region and its statistics are enhanced by the τ-promotion method. Events containing multiple jets without any leptonic objects are estimated with a data-driven ABCD-method. In this analysis, no significant deviation from the expected Standard Model background was observed in data. The signal+background hypothesis was tested in model-dependent simultaneous fits of signal and control regions. For stau masses between 120 GeV and 390 GeV for a massless neutralino LSP and neutralino LSPs up to 140 GeV the signal+background hypothesis could be rejected at 95%. The signal+background hypothesis was also tested for a signal consisting solely of $\tilde{\tau}_L$ pair production. This hypothesis could be rejected between 150 GeV and 320 GeV with neutralino LSP masses up to 70 GeV at 95% CL.

The next aim was to extend the sensitivity of the first stau analysis in multiple directions: Improvements in the lower stau mass phase-space were necessary to come within reach of the the LEP results (which could exclude $\tilde{\tau}_R$ masses up to 90 GeV), while also pushing the sensitivity to even larger stau masses. To achieve this, two boosted decision trees were optimized and trained to deliver a new variable with which the stau signal could be effectively separated from most background events. Again, one BDT was dedicated to the discrimination of signal models with lower stau masses from background (LowMassBDT), while the other was designed especially for higher stau masses (HighMassBDT). Based on the LowMassBDT, four bins were defined that were used in a shape fit, as well as one region used particularly for the model-independent analysis. Another signal region was defined at high scores of the HighMassBDT. The W+jets and Top backgrounds were normalised in dedicated control regions. Again, a data-driven ABCD-method was used to estimate the contribution from multijet events. So far, no comparison with observed data could be made. However, if the background estimate in the signal regions is accurate, the exclusion sensitivity could be extended down to stau masses of 80 GeV and up to 430 GeV. Similarly, the previous limits on $\tilde{\tau}_L$ separately could potentially be extended to cover stau masses of 100 GeV to 350 GeV with neutralino LSPs up to 80 GeV. Looking at the expected limits, a few stau mass points between 120 GeV and 230 GeV and nearly massless neutralino might even be excluded for separate $\tilde{\tau}_R$ pair production. In the absence of any deviation between observed data and expected Standard Model contribution in the

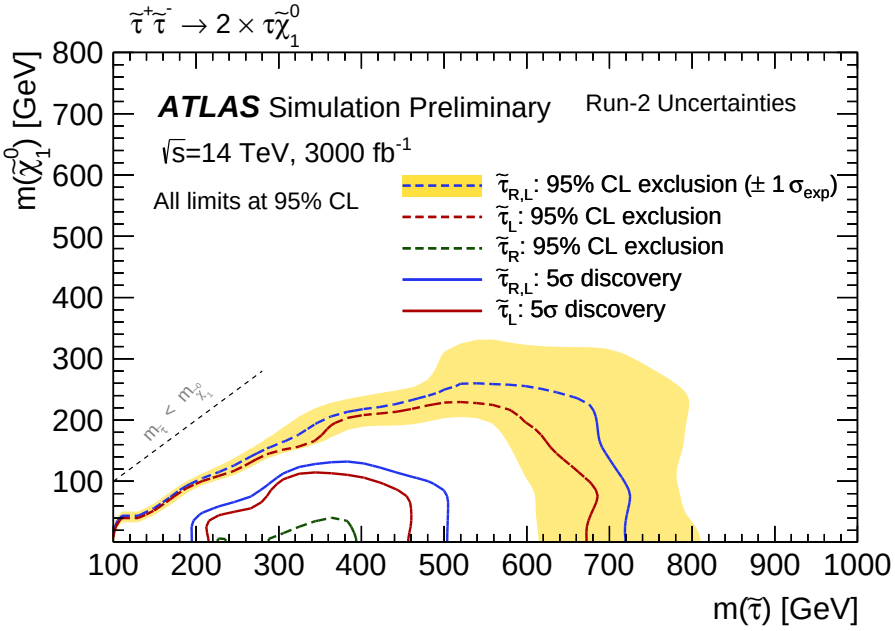


Figure 8.1.: Expected exclusion and discovery sensitivity for stau pair production with high luminosity LHC. The figure is taken from [126].

signal regions, this could result in the first exclusion of $\tilde{\tau}_R$ pair production separate from $\tilde{\tau}_L$ pair production since the results from LEP.

Although up to now there are no strong signs for the existence of staus or Supersymmetry in general in any of the analyses so far performed at the LHC, there is still much space for exploration. A large number of searches for new physics using the data collected in run 2 are still in progress and the development of the LHC and ATLAS machinery is still ongoing. After the "long shutdown 2", the LHC will start its third run, collecting an expected $\sim 300\,\text{fb}^{-1}$ of data. Afterwards, the LHC will be upgraded to become the "High-Lumi" LHC, designed to deliver another $\sim 3000\,\text{fb}^{-1}$ of data. For the search for staus (as well as for every other search for processes with small cross sections), this additional amount of data will increase the potential for discovery (or else exclusion of models) notably, despite new challenges e.g. due to larger pile-up. A study on the prospects for the searches for staus and electroweak gauginos with ATLAS during the High-Lumi LHC shows that not only will the exclusion power be pushed to stau masses as large as 720 GeV, but there will also be the possibility to reach discovery sensitivity for stau masses between 200 GeV and 500 GeV, as can be seen in Fig. 8.1 [126].

Finally, it is important to note that the search for new particles in the context of simplified supersymmetric models is only one part of the overall effort of searching for new physics. There are also other ways like performing pMSSM parameter scans or studies on dark matter, extended Higgs sectors or exotic detector signatures. Hints for new physics might also come from precision measurements of Standard Model parameters. As there is such a large variety of searches and measurements, and the existence of physics beyond the Standard Model is a necessity from both observations and theoretical arguments, there is bound to be a discovery of some kind in the future.

A. Stau Pair Production Cross Sections

Tab. A.1 and Tab. A.2 show the expected production cross sections of stau pairs at proton-proton collisions with a center-of-mass energy of 13 TeV for all considered stau masses at next-to-leading order in perturbation theory.

$m_{\tilde{\tau}}$ [GeV]	Eigenstate	Cross Section [fb]	Relative Uncertainty
80	$\tilde{\tau}_{\mathrm{L}}$	591.02	0.018089
	$\tilde{\tau}_{\mathrm{R}}$	210.37	0.021393
100	$\tilde{\tau}_{\mathrm{L}}$	267.93	0.016790
	$\tilde{\tau}_{\mathrm{R}}$	97.790	0.020012
120	$\tilde{\tau}_{\mathrm{L}}$	140.65	0.016177
	$\tilde{\tau}_{\mathrm{R}}$	52.141	0.020213
140	$\tilde{\tau}_{\mathrm{L}}$	81.153	0.015893
	$\tilde{\tau}_{\mathrm{R}}$	30.404	0.021200
160	$\tilde{\tau}_{\mathrm{L}}$	50.053	0.016417
	$\tilde{\tau}_{\mathrm{R}}$	18.909	0.021946
180	$\tilde{\tau}_{\mathrm{L}}$	32.507	0.017659
	$\tilde{\tau}_{\mathrm{R}}$	12.344	0.022037
200	$\tilde{\tau}_{\mathrm{L}}$	21.942	0.019183
	$\tilde{\tau}_{\mathrm{R}}$	8.3720	0.021821
220	$\tilde{\tau}_{\mathrm{L}}$	15.291	0.020081
	$\tilde{\tau}_{\mathrm{R}}$	5.8578	0.021121
240	$\tilde{\tau}_{\mathrm{L}}$	10.939	0.020338
	$\tilde{\tau}_{\mathrm{R}}$	4.2038	0.021489
260	$\tilde{\tau}_{\mathrm{L}}$	7.9974	0.020549
	$\tilde{\tau}_{\mathrm{R}}$	3.0821	0.022107
280	$\tilde{\tau}_{\mathrm{L}}$	5.9566	0.020931
	$\tilde{\tau}_{\mathrm{R}}$	2.3019	0.022214
300	$\tilde{\tau}_{\mathrm{L}}$	4.5083	0.021737
	$\tilde{\tau}_{\mathrm{R}}$	1.7459	0.022799
320	$\tilde{\tau}_{\mathrm{L}}$	3.4597	0.022526
	$\tilde{\tau}_{\mathrm{R}}$	1.3424	0.022795
340	$\tilde{\tau}_{\mathrm{L}}$	2.6873	0.024240
	$\tilde{\tau}_{\mathrm{R}}$	1.0445	0.022937
360	$\tilde{\tau}_{\mathrm{L}}$	2.1099	0.026029
	$\tilde{\tau}_{\mathrm{R}}$	0.8215	0.024637
380	$\tilde{\tau}_{\mathrm{L}}$	1.6724	0.027226
	$\tilde{\tau}_{\mathrm{R}}$	0.6522	0.025985
400	$\tilde{\tau}_{\mathrm{L}}$	1.3370	0.029027
	$\tilde{\tau}_{\mathrm{R}}$	0.5222	0.027477
440	$\tilde{\tau}_{\mathrm{L}}$	0.8738	0.031916
	$\tilde{\tau}_{\mathrm{R}}$	0.3422	0.028306

Table A.1.: Production cross section of stau pairs in dependence on the stau mass and the stau eigenstate. The cross sections have been calculated with the Resummino package. [40–44]

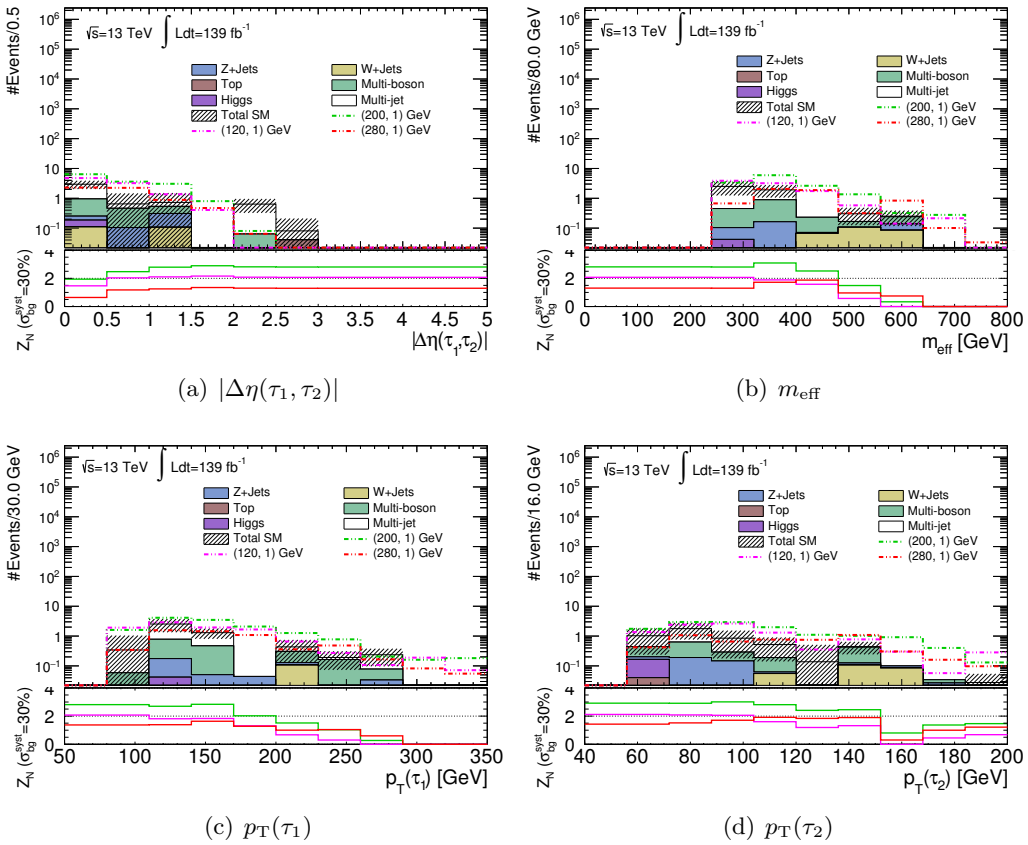
$m_{\tilde{\tau}}$ [GeV]	Eigenstate	Cross Section [fb]	Relative Uncertainty
480	$\tilde{\tau}_{\mathrm{L}}$	0.5856	0.032819
	$\tilde{\tau}_{\mathrm{R}}$	0.2300	0.030713
520	$\tilde{\tau}_{\mathrm{L}}$	0.4010	0.034018
	$\tilde{\tau}_{\mathrm{R}}$	0.1579	0.031536
560	$\tilde{\tau}_{\mathrm{L}}$	0.2796	0.036090
	$\tilde{\tau}_{\mathrm{R}}$	0.1103	0.033347

Table A.2.: Production cross section of stau pairs in dependence on the stau mass and the stau eigenstate. The cross sections have been calculated with the Resummino package. [40–44]

B. First Search: N-1 Plots

The following plots show the kinematic distributions for the signal regions SR-lowMass and SR-highMass described in Sec. 6.2.2. All of the cuts listed in Tab. 6.3 are applied, except the cut on the showed variable if there is any ("N-1"). The cut on the plotted variable and its direction is indicated by a red arrow. Only cuts derived in the signal region optimization process are shown, in contrast to the preselection and trigger plateau cuts. The lower panel shows the significance a cut on the respective value on the x-axis would result in. The considered value on the x-axis is taken to be an upper or lower cut, depending on the type of cut that is looked for in the signal region optimization process.

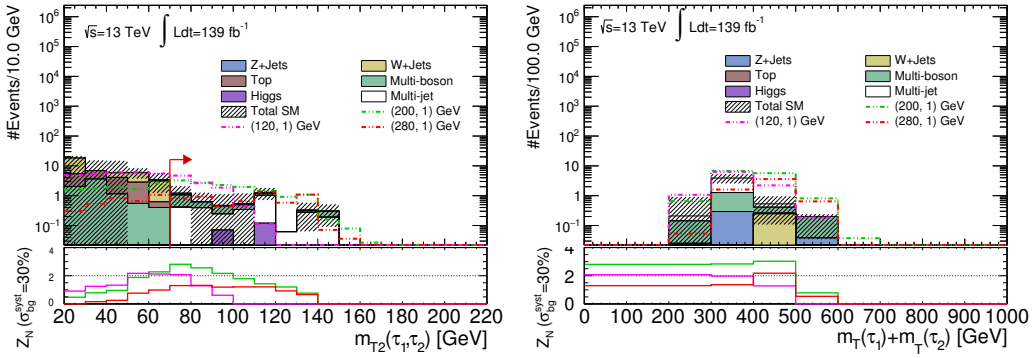
B.1. N-1 Plots for SR-lowMass



(a) $|\Delta\eta(\tau_1, \tau_2)|$ (b) m_{eff}

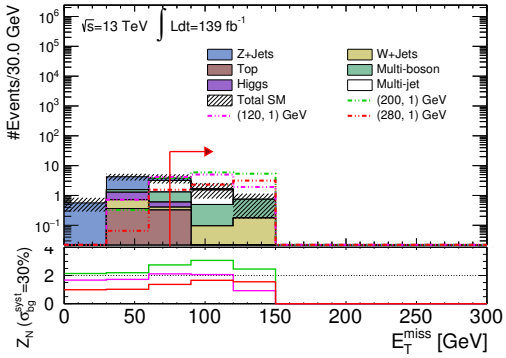
(c) $p_{\text{T}}(\tau_1)$ (d) $p_{\text{T}}(\tau_2)$

Figure B.1.: Kinematic distributions in SR-lowMass. All cuts are applied except the cut on the plotted variable (if any), which is indicated by a red arrow. The lower panel shows the significance in dependence of the shown variable. A flat uncertainty of 30% is considered in addition to the statistical uncertainties of the background, to account for systematic uncertainties.

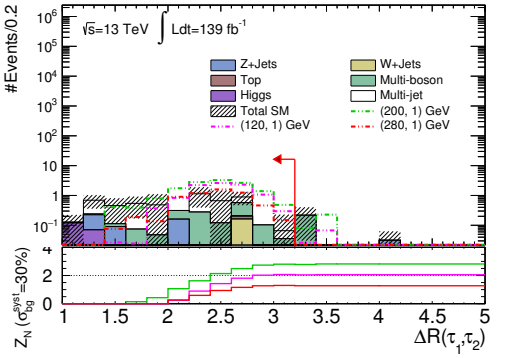


(a) $m_{\mathrm{T2}}(\tau_1, \tau_2)$

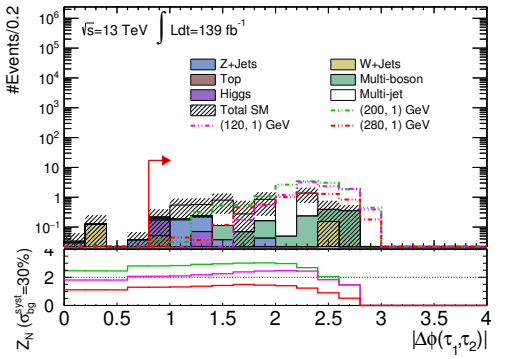
(b) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



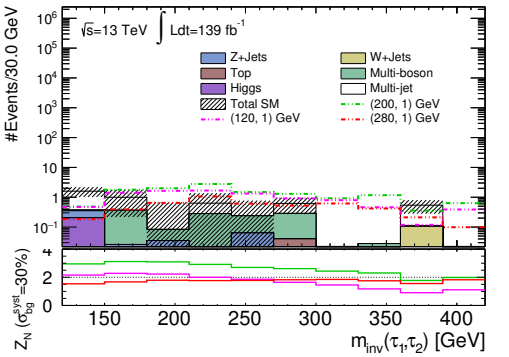
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$



(d) $\Delta R(\tau_1, \tau_2)$



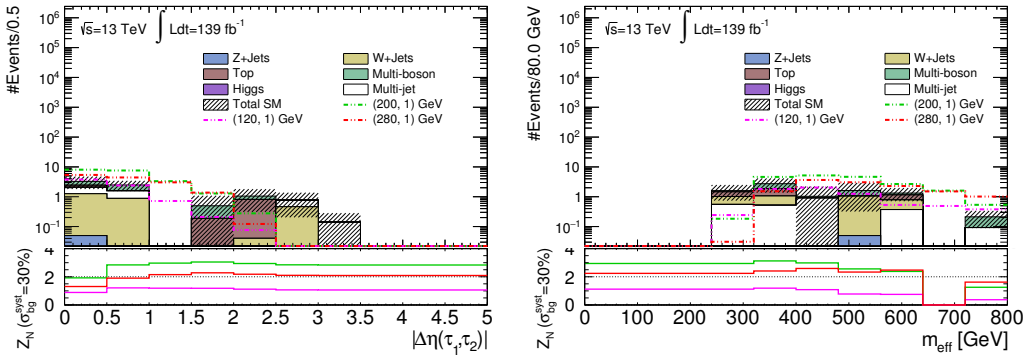
(e) $|\Delta\phi(\tau_1, \tau_2)|$



(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

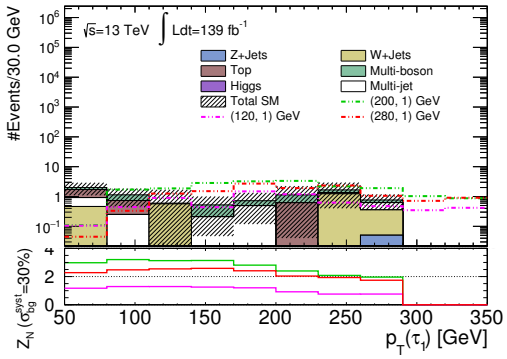
Figure B.2.: Kinematic distributions in SR-lowMass. All cuts are applied except the cut on the plotted variable (if any), which is indicated by a red arrow. The lower panel shows the significance in dependence of the shown variable. A flat uncertainty of 30% is considered in addition to the statistical uncertainties of the background, to account for systematic uncertainties.

B.2. N-1 Plots for SR-highMass

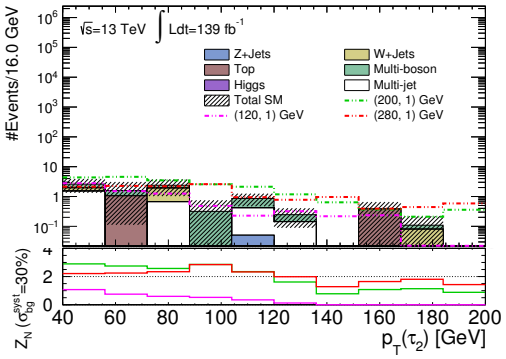


(a) $|\Delta\eta(\tau_1, \tau_2)|$

(b) m_{eff}

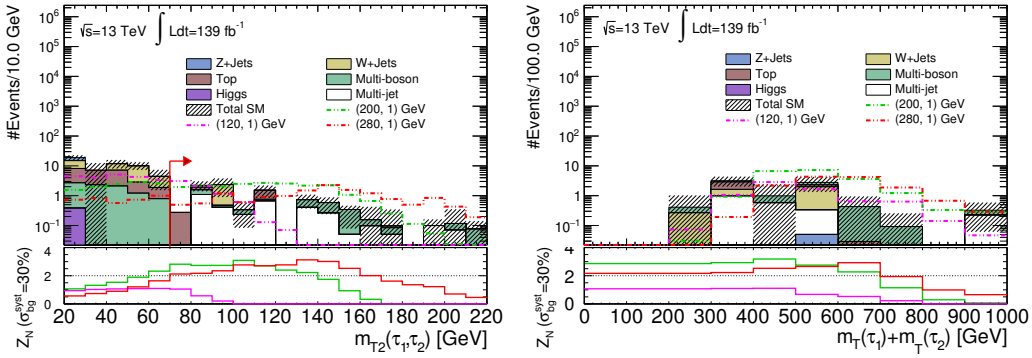


(c) $p_{\text{T}}(\tau_1)$



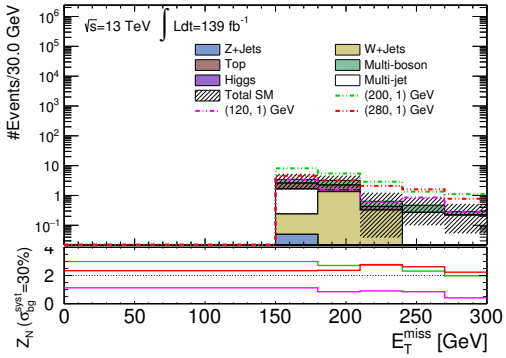
(d) $p_{\text{T}}(\tau_2)$

Figure B.3.: Kinematic distributions in SR-highMass. All cuts are applied except the cut on the plotted variable (if any), which is indicated by a red arrow. The lower panel shows the significance in dependence of the shown variable. A flat uncertainty of 30% is considered in addition to the statistical uncertainties of the background, to account for systematic uncertainties.

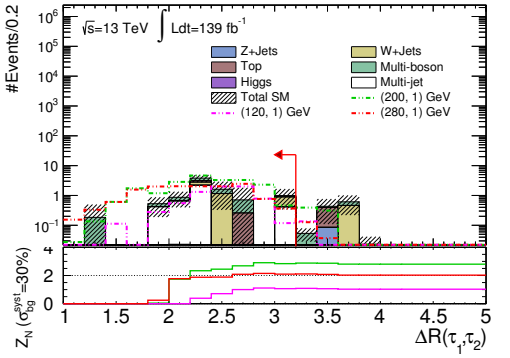


(a) $m_{\mathrm{T2}}(\tau_1, \tau_2)$

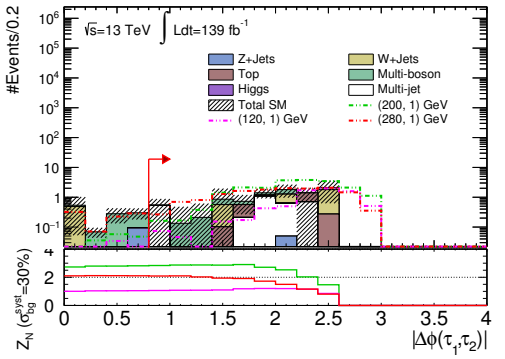
(b) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



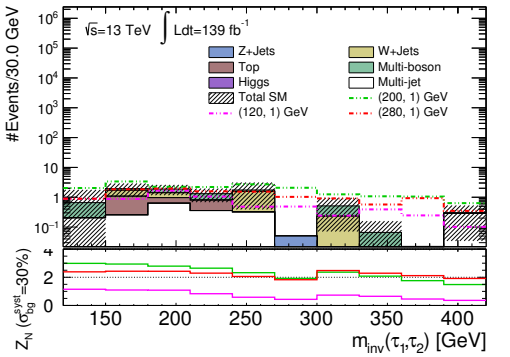
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$



(d) $\Delta R(\tau_1, \tau_2)$



(e) $|\Delta\phi(\tau_1, \tau_2)|$



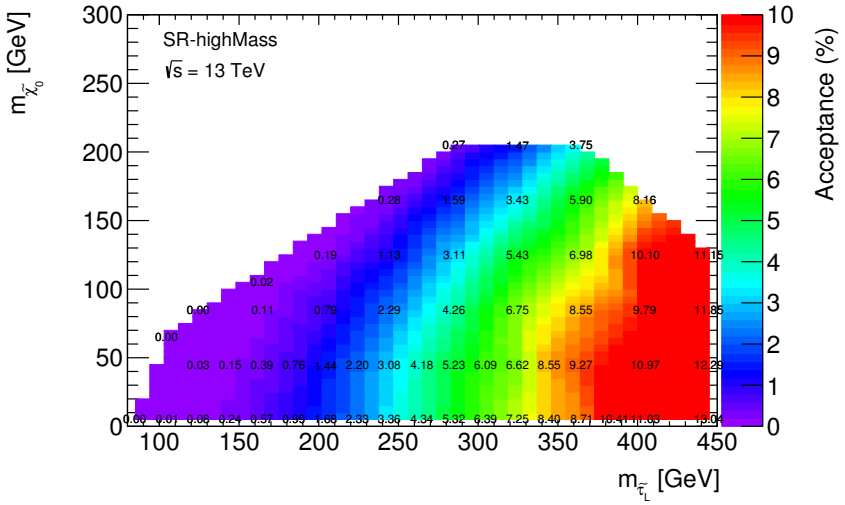
(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

Figure B.4.: Kinematic distributions in SR-highMass. All cuts are applied except the cut on the plotted variable (if any), which is indicated by a red arrow. The lower panel shows the significance in dependence of the shown variable. A flat value of 30% is considered in addition to the statistical uncertainties of the background, to account for systematic uncertainties.

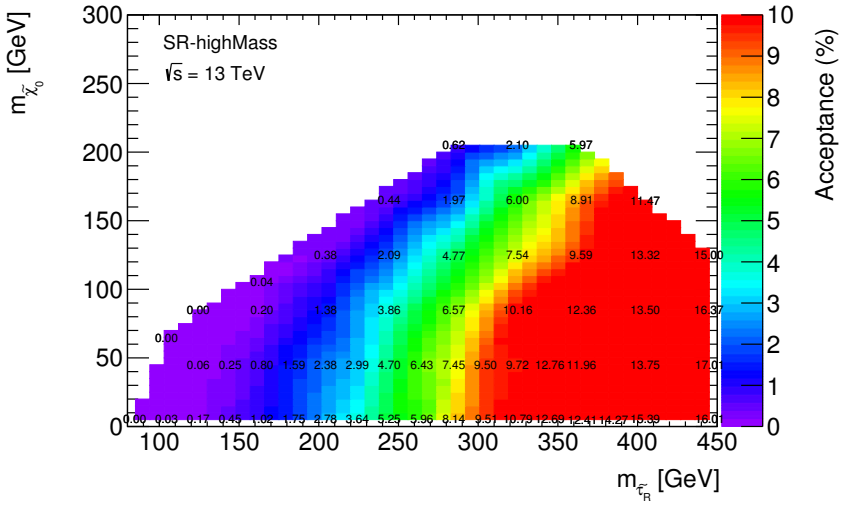
C. First Search: Additional Material on Acceptance and Efficiency in SRs

In this section, additional results concerning the separate production scenarios for the stau eigenstates for the acceptance and efficiency studies in Sec. 6.7 are shown.

Acceptance

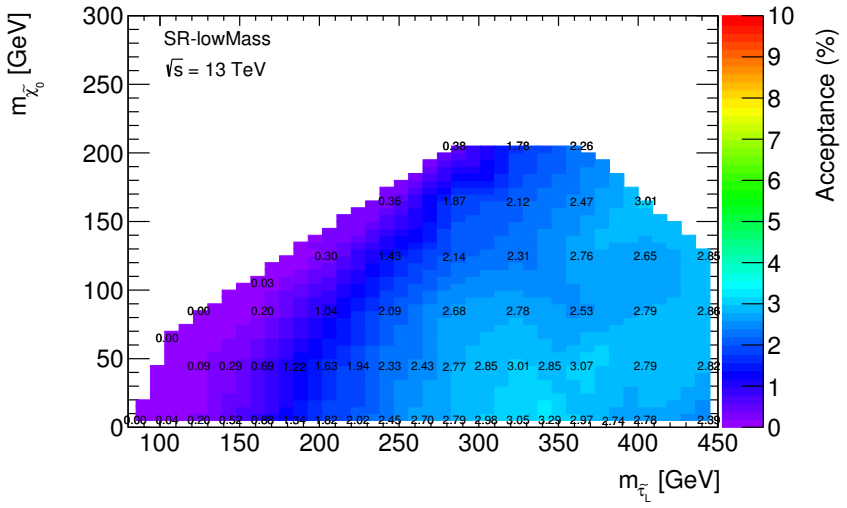


(a) $\tilde{\tau}_L\tilde{\tau}_L$

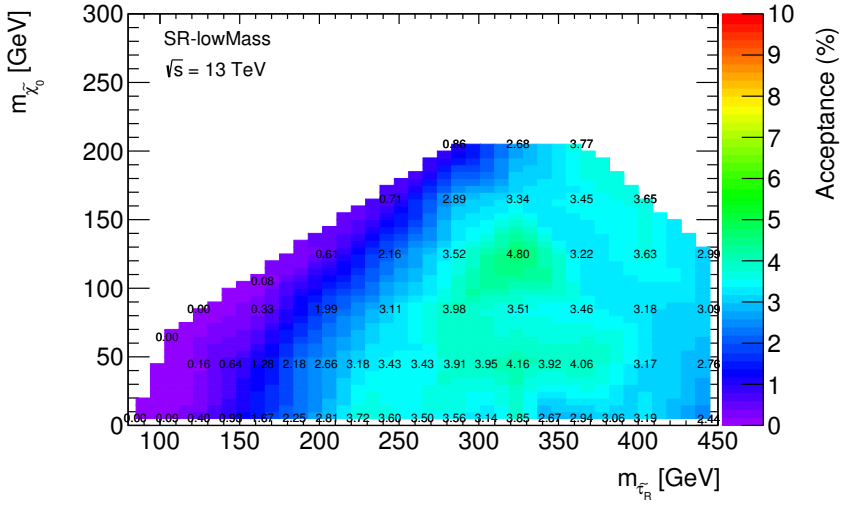


(b) $\tilde{\tau}_R\tilde{\tau}_R$

Figure C.1.: Signal acceptance for different stau and neutralino masses in the highMass signal region for the separate pair production of each of stau eigenstates.



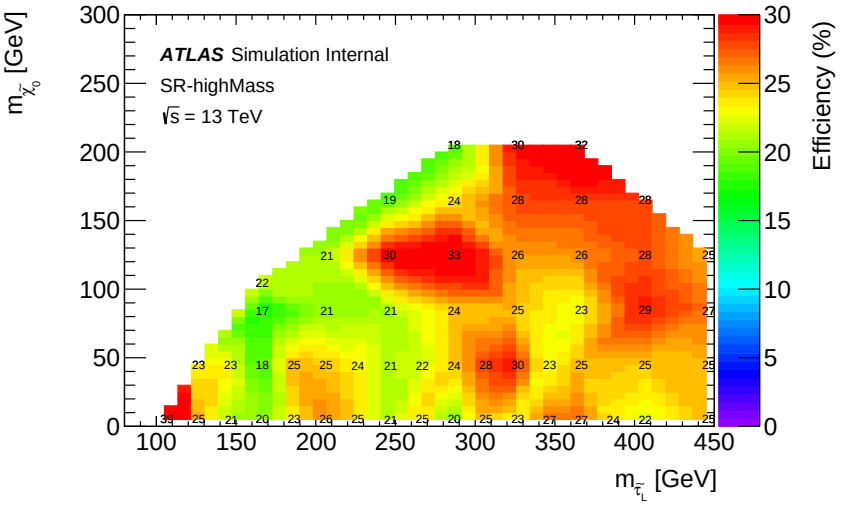
(a) $\tilde{\tau}_L\tilde{\tau}_L$



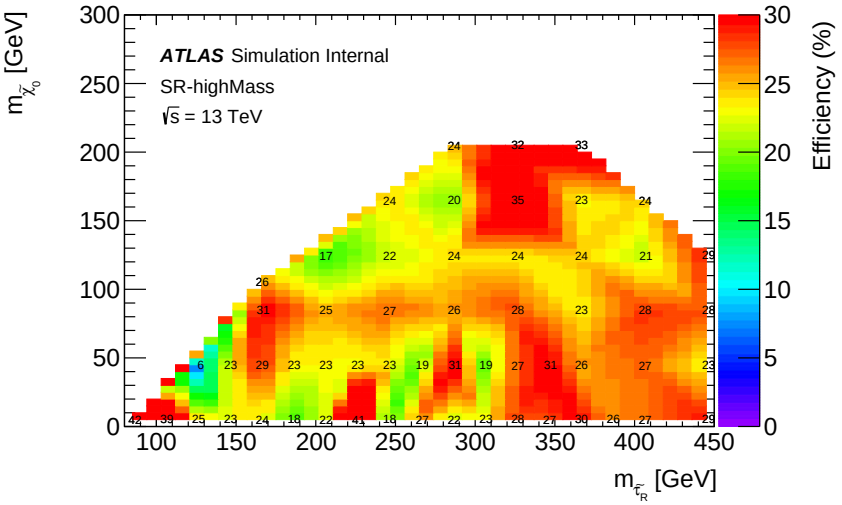
(b) $\tilde{\tau}_R\tilde{\tau}_R$

Figure C.2.: Signal acceptance for different stau and neutralino masses in the lowMass signal region for the separate pair production each of the stau eigenstates.

Efficiency

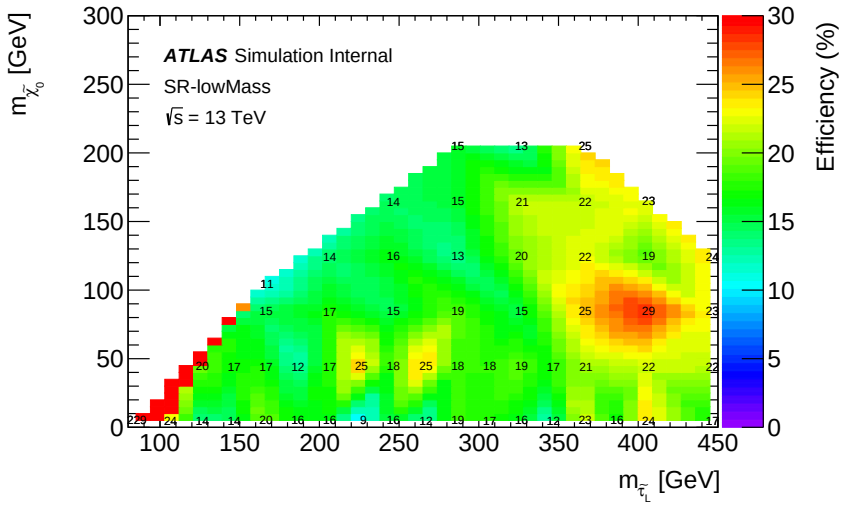


(a) $\tilde{\tau}_L\tilde{\tau}_L$

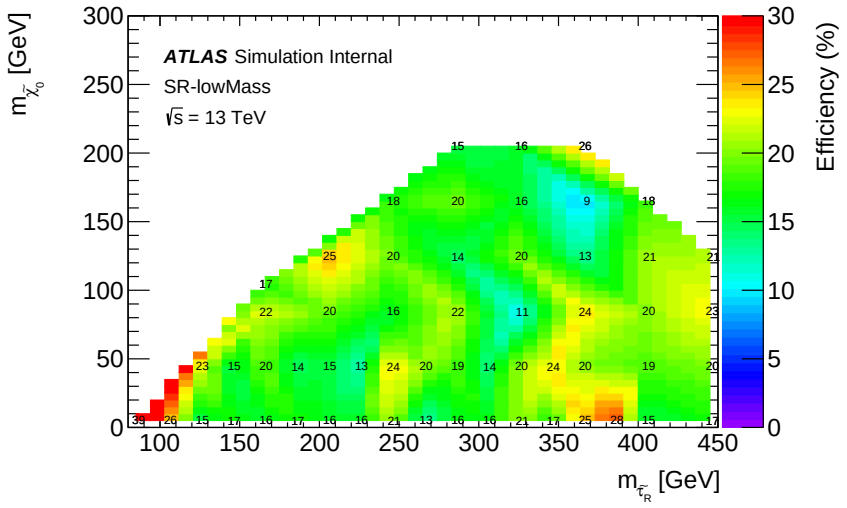


(b) $\tilde{\tau}_R\tilde{\tau}_R$

Figure C.3.: Signal efficiency for different stau and neutralino masses in the highMass signal region for the separate pair production each of the stau eigenstates.



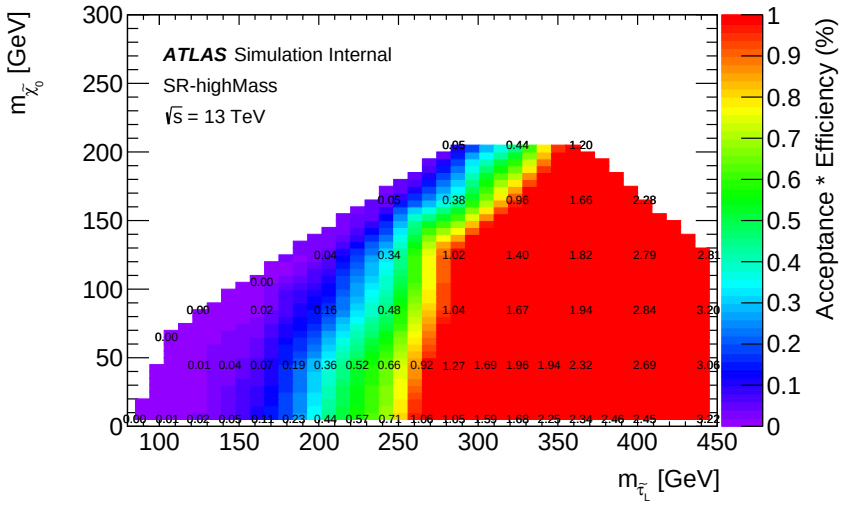
(a) $\tilde{\tau}_L\tilde{\tau}_L$



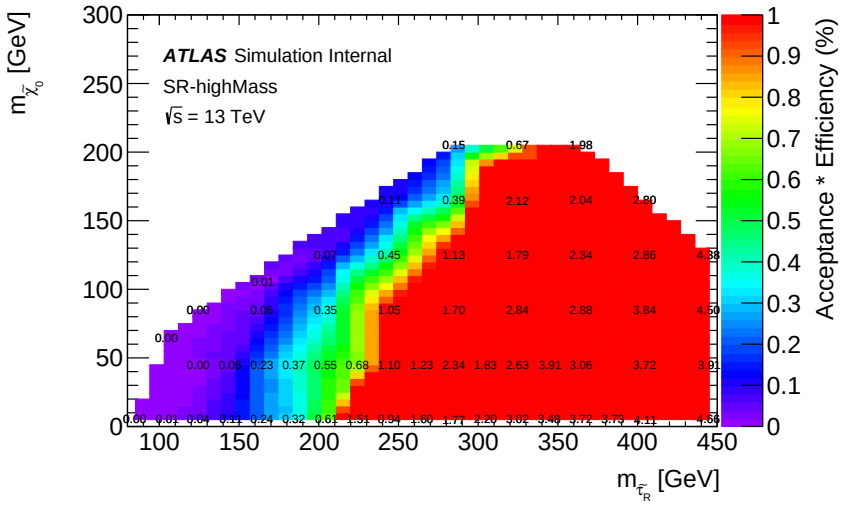
(b) $\tilde{\tau}_R\tilde{\tau}_R$

Figure C.4.: Signal efficiency for different stau and neutralino masses in the lowMass signal region for the separate pair production each of the stau eigenstates.

Acceptance×Efficiency

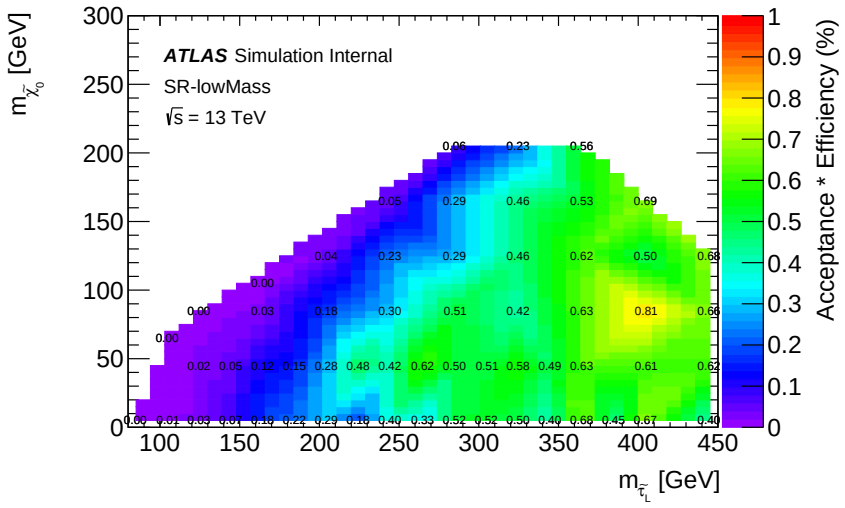


(a) $\tilde{\tau}_L\tilde{\tau}_L$

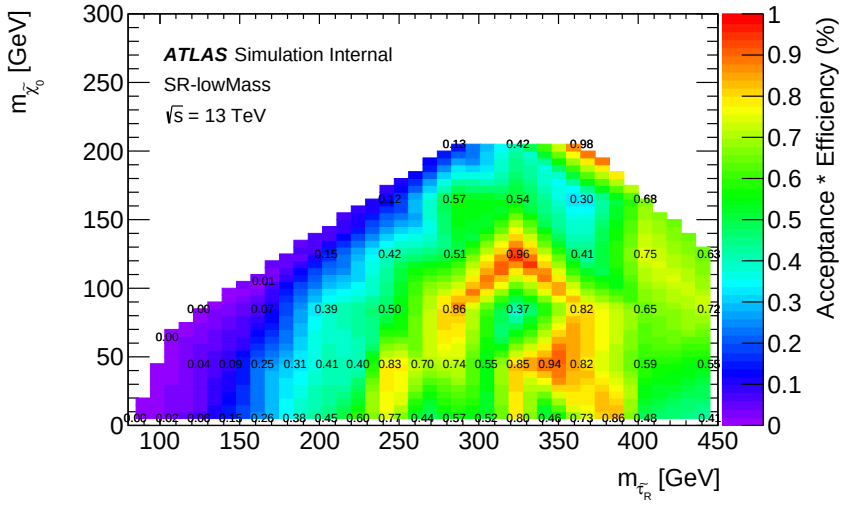


(b) $\tilde{\tau}_R\tilde{\tau}_R$

Figure C.5.: Signal acceptance×efficiency for different stau and neutralino masses in the highMass signal region for the separate pair production each of the stau eigenstates.



(a) $\tilde{\tau}_L\tilde{\tau}_L$



(b) $\tilde{\tau}_R\tilde{\tau}_R$

Figure C.6.: Signal acceptance×efficiency for different stau and neutralino masses in the lowMass signal region for the separate pair production each of the stau eigenstates.

D. Second Search: Additional Material on BDTs

D.1. Input Variable Ranking

The following tables show the importance of the input variables in the form of a ranking. The higher the rank of a variable, the more often it was used during the training. For both BDTs the most important input variables are $m_{\mathrm{T2}}(\tau_1, \tau_2)$ and the sum of the transverse masses of the two τ-leptons.

Rank	Variable
1	$m_{\mathrm{T2}}(\tau_1, \tau_2)$ $[m_{\mathrm{invisible}} = 0\,\mathrm{GeV}]$
2	$m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$
3	$\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$
4	$\Delta\phi(\tau_1, \tau_2)$
5	$E_{\mathrm{T}}^{\mathrm{miss}}$
6	$m_{\mathrm{T2}}(\tau_1, \tau_2)$ $[m_{\mathrm{invisible}} = 40\,\mathrm{GeV}]$
7	$\Delta R(\tau_1, \tau_2)$
8	$\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$
9	$p_{\mathrm{T}}(\tau_2)$
10	$\Delta\eta(\tau_1, \tau_2)$
11	$m_{\mathrm{inv}}(\tau_1, \tau_2)$
12	$p_{\mathrm{T}}(\mathrm{jet}_1)$
13	m_{eff}
14	$p_{\mathrm{T}}(\tau_1)$
15	$p_{\mathrm{T}}(\mathrm{jet}_2)$
16	$N(\mathrm{jets})$
17	$N(b\text{-tags})$

Table D.1.: List of variables used for training the LowMassBDT, ordered by their respective "importance" to the BDT, which is a measure of how often the variable was used during the training.

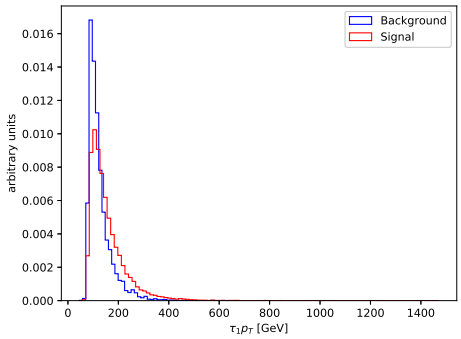
Rank	Variable
1	$m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$
2	$m_{\mathrm{T2}}(\tau_1, \tau_2)$ $[m_{\mathrm{invisible}} = 0\,\mathrm{GeV}]$
3	$\Delta R(\tau_1, \tau_2)$
4	$m_{\mathrm{T2}}(\tau_1, \tau_2)$ $[m_{\mathrm{invisible}} = 40\,\mathrm{GeV}]$
5	$p_{\mathrm{T}}(\mathrm{jet}_1)$
6	$m_{\mathrm{inv}}(\tau_1, \tau_2)$
7	$\Delta\phi(\tau_1, \tau_2)$
8	$N(b\text{-tags})$
9	$p_{\mathrm{T}}(\tau_2)$
10	$E_{\mathrm{T}}^{\mathrm{miss}}$
11	$\Delta\eta(\tau_1, \tau_2)$
12	$p_{\mathrm{T}}(\tau_1)$
13	$\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$
14	$\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$
15	$p_{\mathrm{T}}(\mathrm{jet}_2)$
16	$N(\mathrm{jets})$
17	m_{eff}

Table D.2.: List of variables used for training the HighMassBDT, ordered by their respective "importance" to the BDT, which is a measure of how often the variable was used during the training.

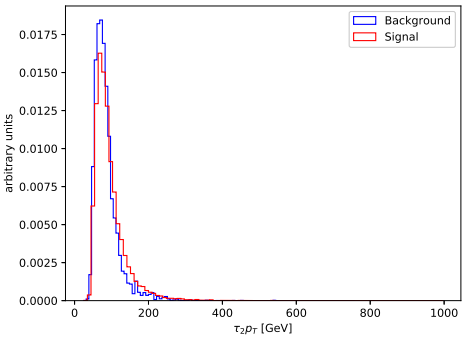
D.2. BDT Input Variable Shape Plots

In the following plots, the discrimination power of the individual input variables to the BDTs is shown. The distributions are normalized and combine all signal models used for training together (red) as well as all background samples (blue). In the plots for both pre-training selections for $\Delta\phi(\tau_1, \tau_2)$ and $\Delta R(\tau_1, \tau_2$ a bump is visible in the background distribution which is not present in the signal distribution. This bump is predominantly caused by Z+jets, which can also be seen in the $\Delta R(\tau_1, \tau_2)$ plot in Fig. E.10.

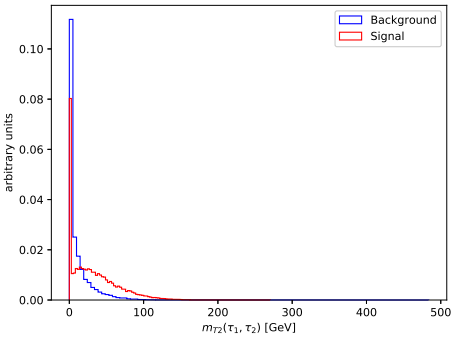
D.2.1. Shape Plots for LowMassBDT Input Variables



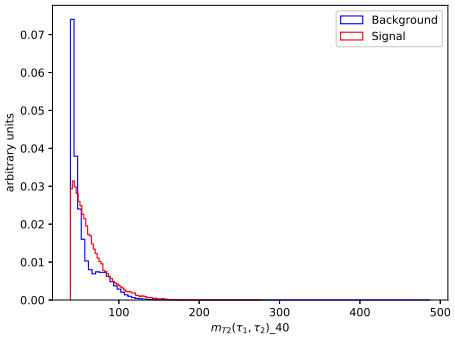
(a) $p_{\mathrm{T}}(\tau_1)$



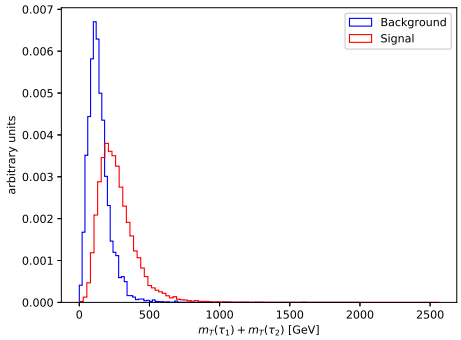
(b) $p_{\mathrm{T}}(\tau_2)$



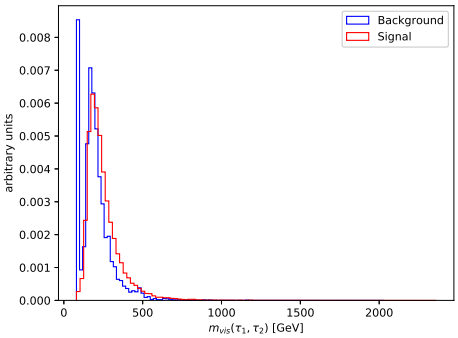
(c) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ $(m_\chi = 0\,\mathrm{GeV})$



(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ $(m_\chi = 40\,\mathrm{GeV})$

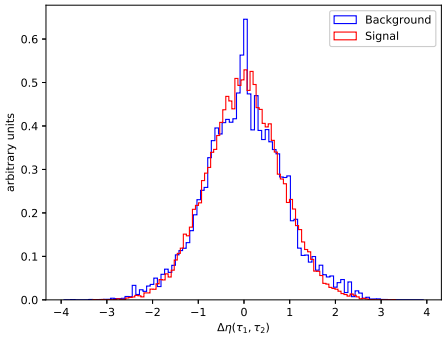


(e) $m_{\mathrm{T}}(\tau_1)$+$m_{\mathrm{T}}(\tau_2)$

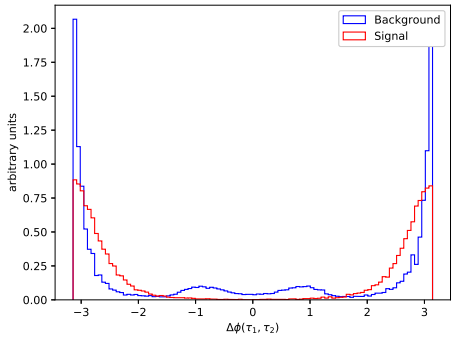


(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

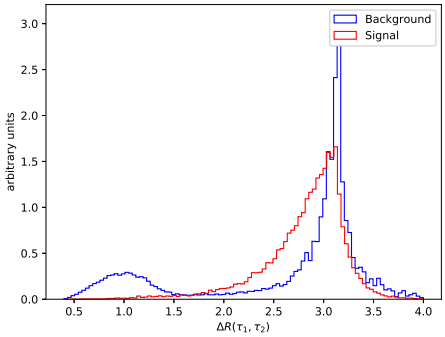
Figure D.1.: Shape plots for the LowMassBDT input variables.



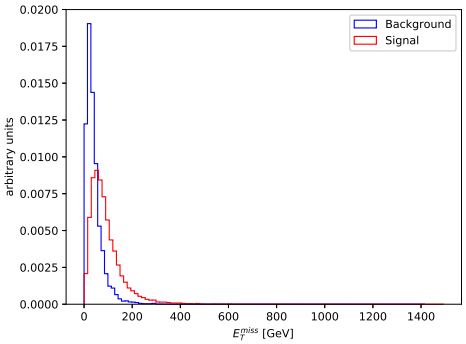
(a) $\Delta\eta(\tau_1, \tau_2)$



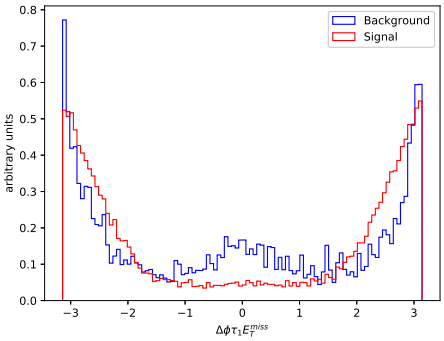
(b) $\Delta\phi(\tau_1, \tau_2)$



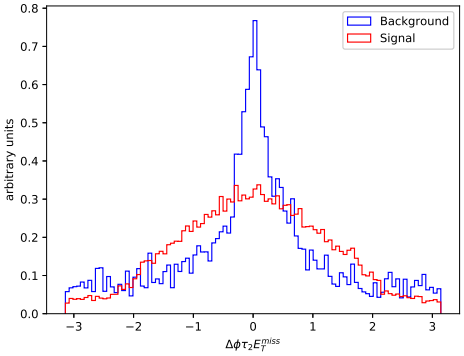
(c) $\Delta R(\tau_1, \tau_2)$



(d) $E_{\mathrm{T}}^{\mathrm{miss}}$

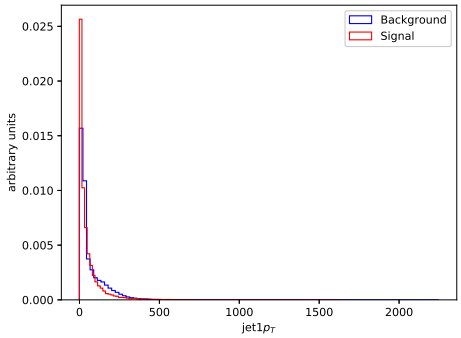


(e) $\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$

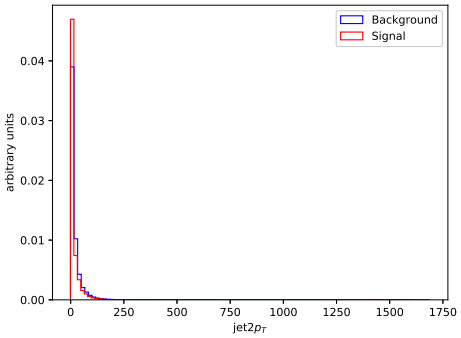


(f) $\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$

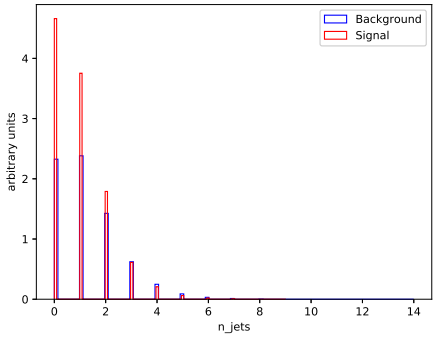
Figure D.2.: Shape plots for the LowMassBDT input variables.



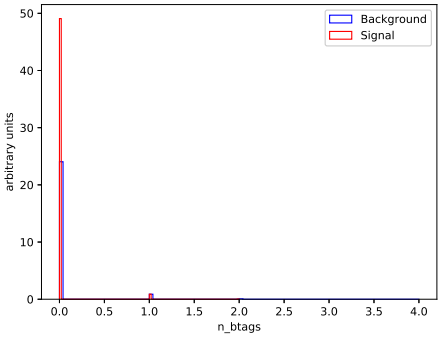
(a) $p_{\mathrm{T}}(\mathrm{jet}_1)$



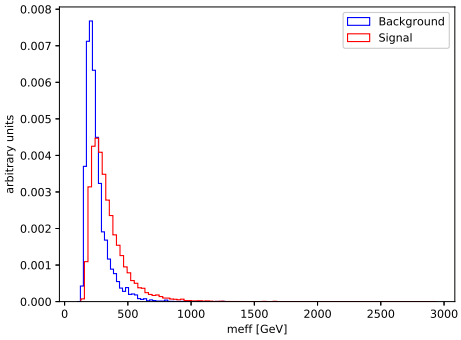
(b) $p_{\mathrm{T}}(\mathrm{jet}_2)$



(c) $N(\mathrm{jets})$



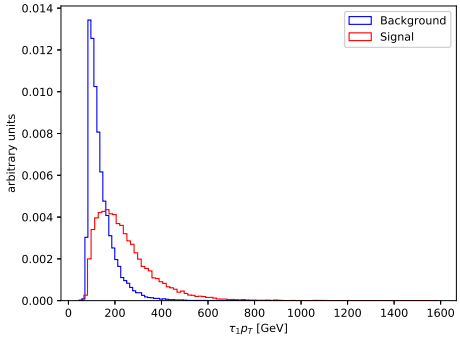
(d) $N(\mathrm{b\text{-}tags})$



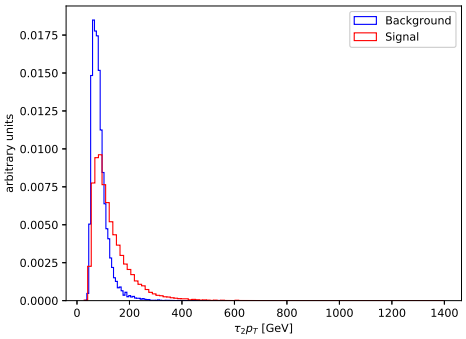
(e) m_{eff}

Figure D.3.: Shape plots for the LowMassBDT input variables.

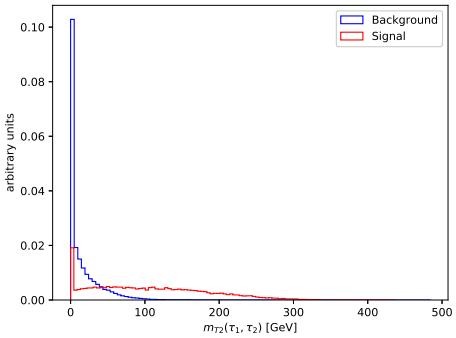
D.2.2. Shape Plots for HighMassBDT Input Variables



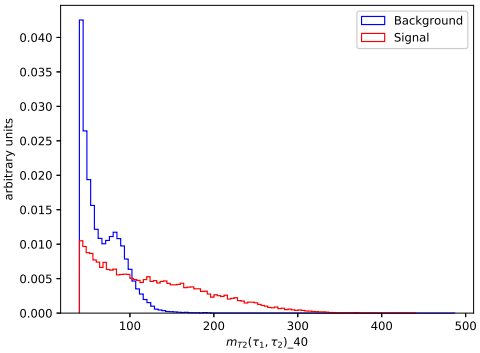
(a) $p_{\mathrm{T}}(\tau_1)$



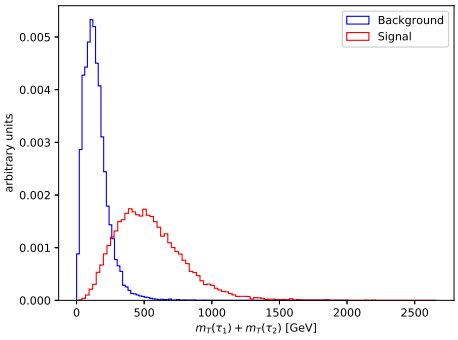
(b) $p_{\mathrm{T}}(\tau_2)$



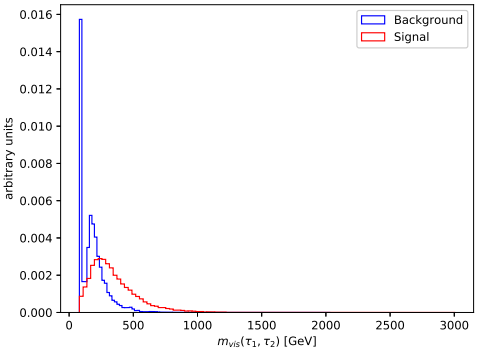
(c) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ ($m_\chi = 0\,\mathrm{GeV}$)



(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ ($m_\chi = 40\,\mathrm{GeV}$)

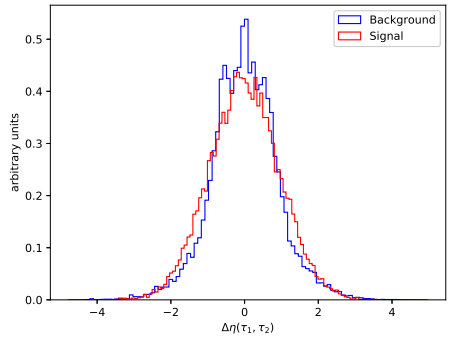


(e) $m_{\mathrm{T}}(\tau_1)$+$m_{\mathrm{T}}(\tau_2)$

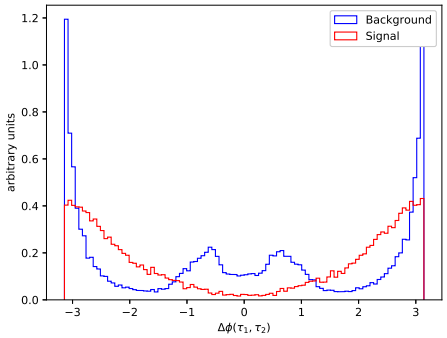


(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

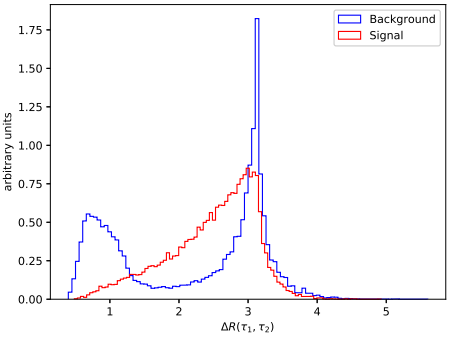
Figure D.4.: Shape plots for the HighMassBDT input variables.



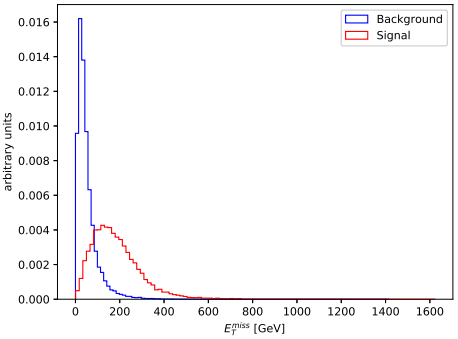
(a) $\Delta\eta(\tau_1, \tau_2)$



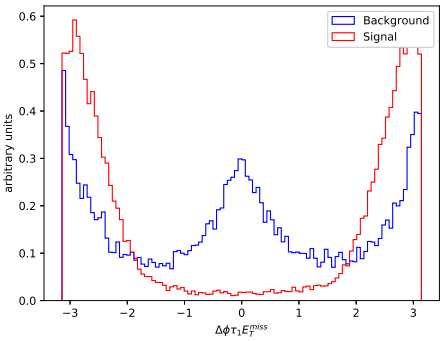
(b) $\Delta\phi(\tau_1, \tau_2)$



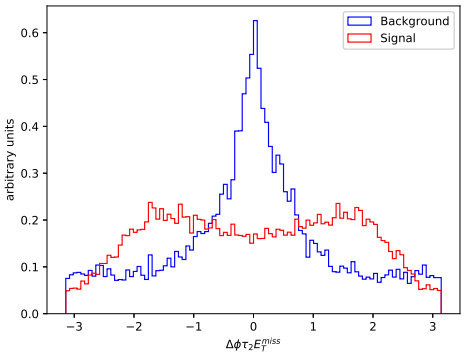
(c) $\Delta R(\tau_1, \tau_2)$



(d) $E_{\mathrm{T}}^{\mathrm{miss}}$

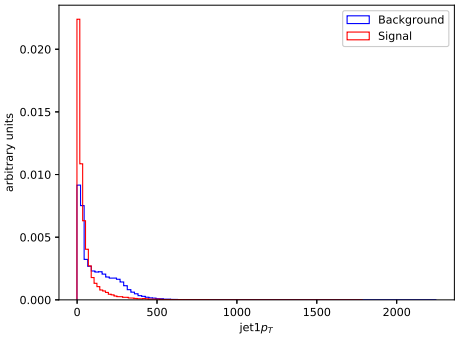


(e) $\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$

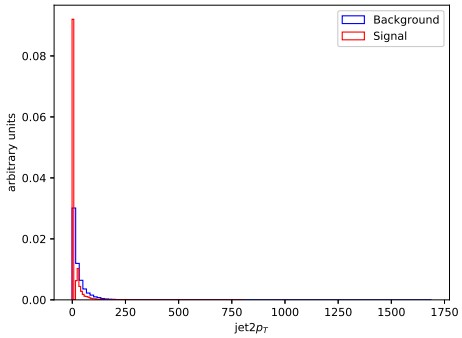


(f) $\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$

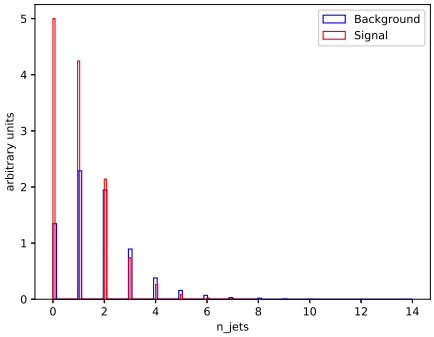
Figure D.5.: Shape plots for the HighMassBDT input variables.



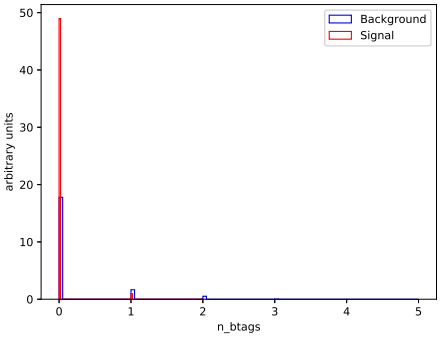
(a) $p_{\mathrm{T}}(\mathrm{jet}_1)$



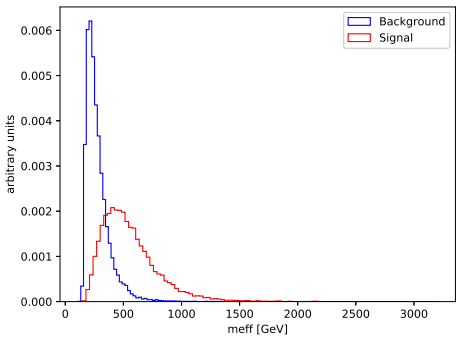
(b) $p_{\mathrm{T}}(\mathrm{jet}_2)$



(c) $N(\mathrm{jets})$



(d) $N(\text{b-tags})$

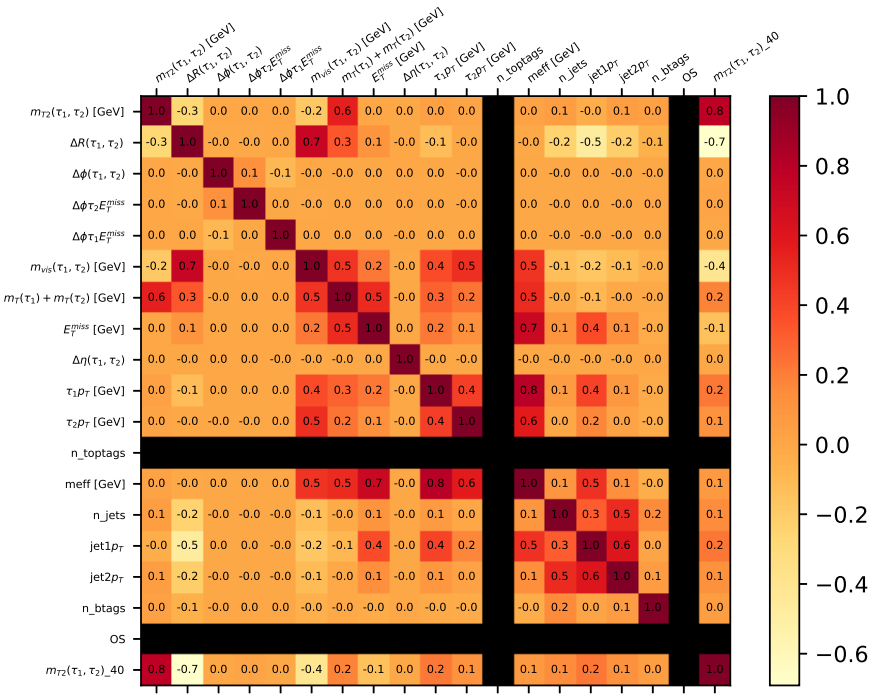


(e) m_{eff}

Figure D.6.: Shape plots for the HighMassBDT input variables.

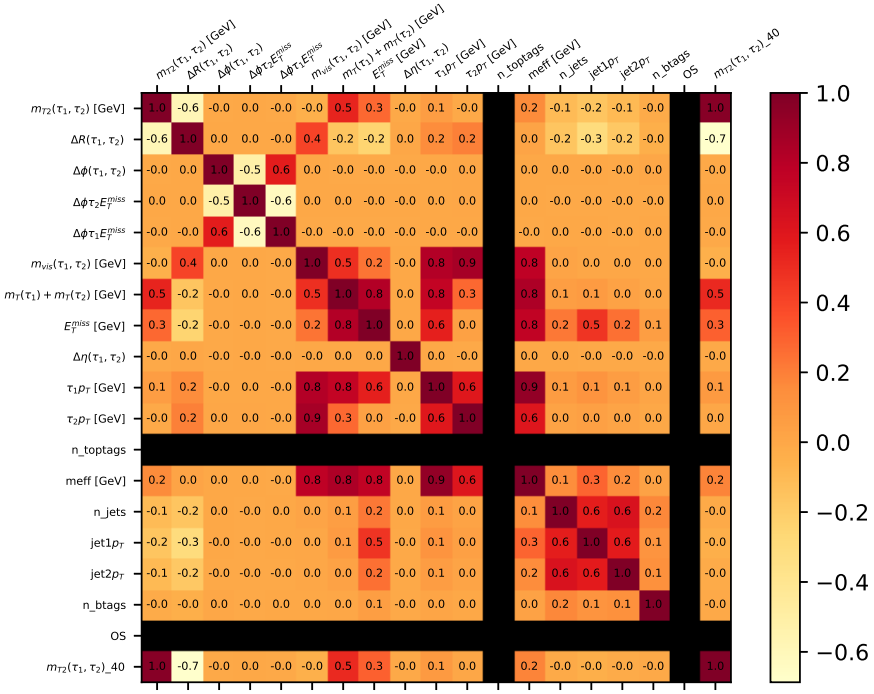
D.3. Linear Correlation of BDT Input Variables

The linear correlation parameters between the input variables for training are shown for both the signal and the background dataset for the LowMassBDT and the HighMassBDT.



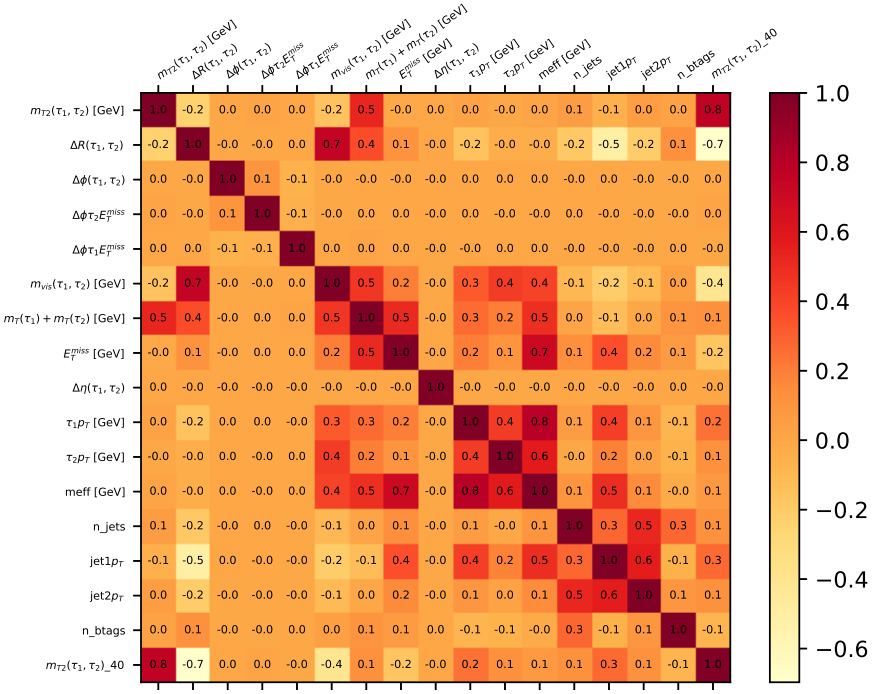
(a) LowMassBDT Signal

Figure D.7.: Linear correlation of LowMassBDT input variables for the signal dataset. The black columns and rows for the LowMassBDT matrices correspond to variables that were included in the training originally but were set to constant values in the course of the BDT optimization.



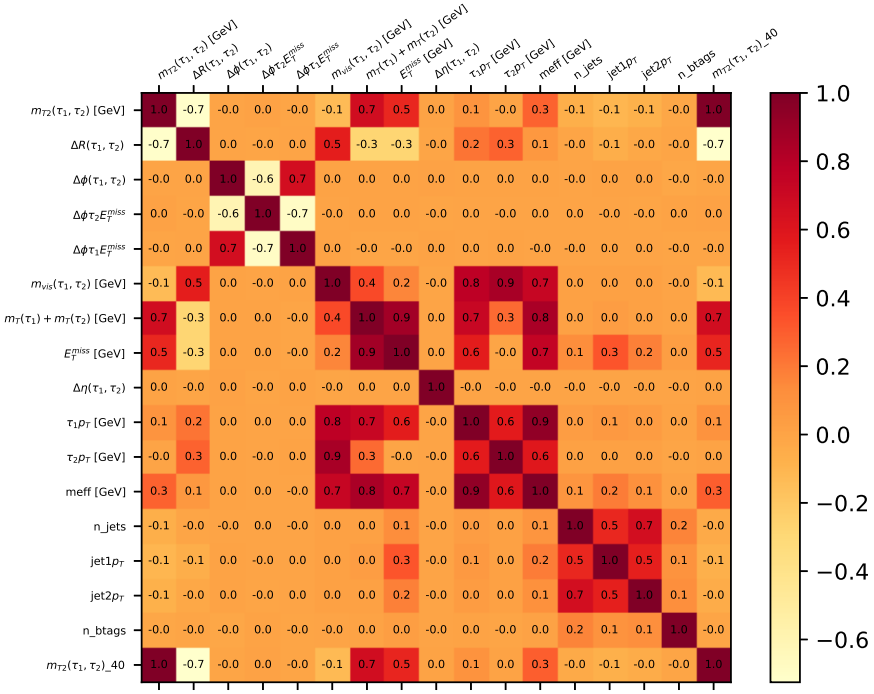
(a) LowMassBDT Background

Figure D.8.: Linear correlation of LowMassBDT input variables for the background dataset. The black columns and rows for the LowMassBDT matrices correspond to variables that were included in the training originally but were set to constant values in the course of the BDT optimization.



(a) HighMassBDT Signal

Figure D.9.: Linear correlation of HighMassBDT input variables for the signal dataset.



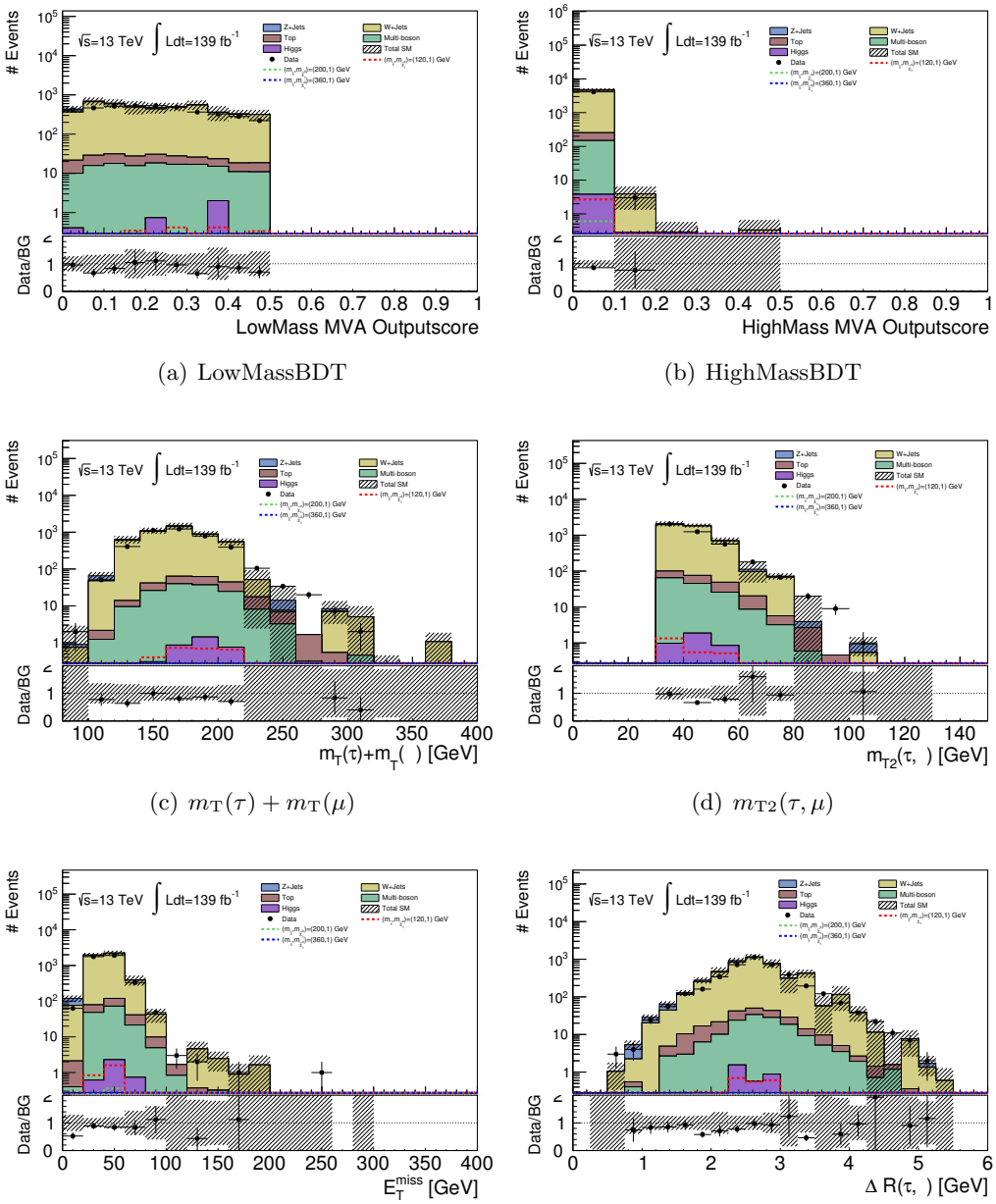
(a) HighMassBDT Background

Figure D.10.: Linear correlation of HighMassBDT input variables for the background dataset.

E. Second Search: Additional Material on Background Estimation

In the following, several kinematic distributions are shown for all used control and validation regions. Sec. E.5 contains additional material on the ABCD method used for multijet estimation and its results.

E.1. W+jets Estimation



(a) LowMassBDT

(b) HighMassBDT

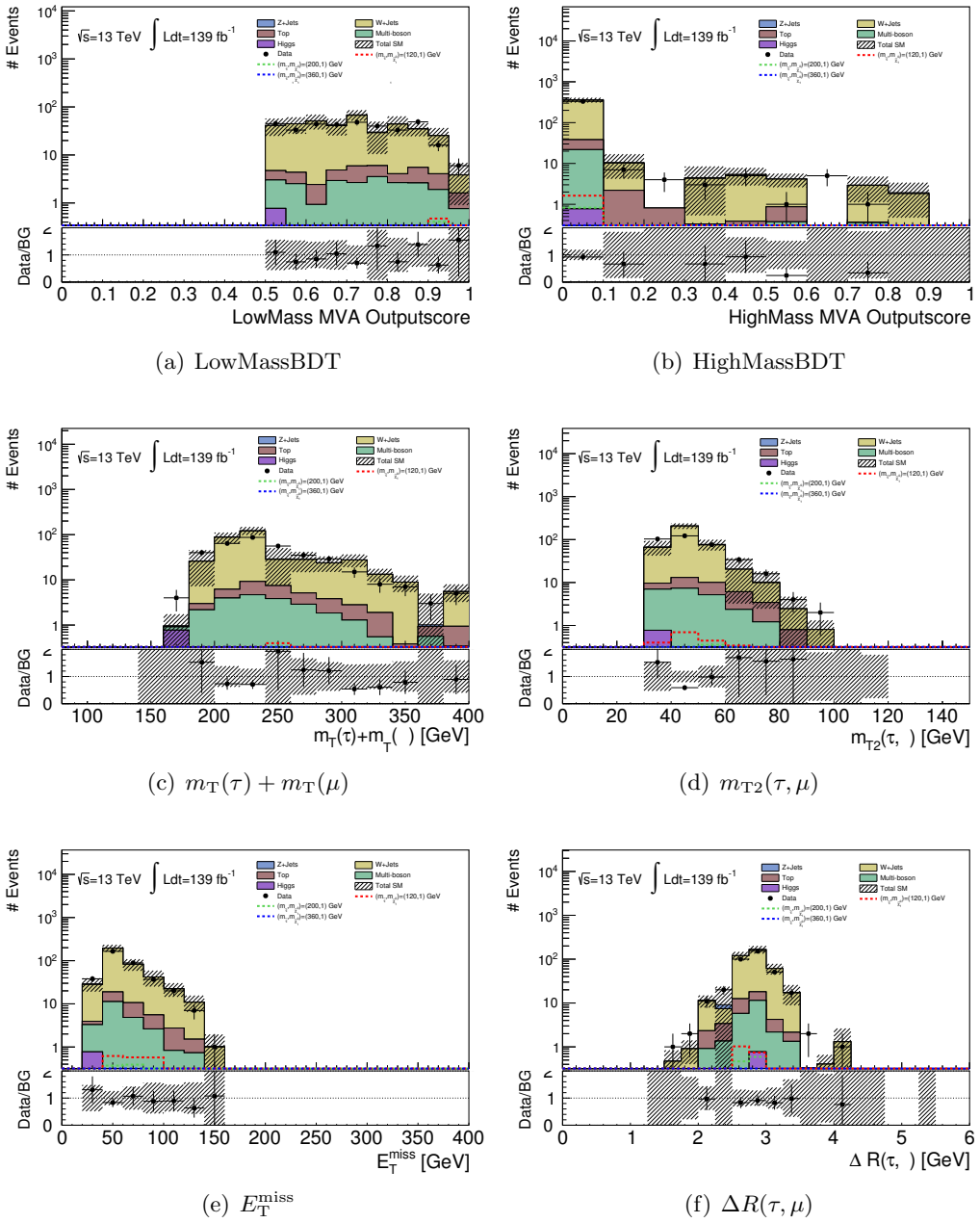
(c) $m_{\mathrm{T}}(\tau) + m_{\mathrm{T}}(\mu)$

(d) $m_{\mathrm{T2}}(\tau, \mu)$

(e) $E_{\mathrm{T}}^{\mathrm{miss}}$

(f) $\Delta R(\tau, \mu)$

Figure E.1.: Pre-fit kinematic distributions for W-CR. For W-CR, W-VR lowMass and W-VR highMass events containing one signal τ-lepton and one isolated muon are selected. The W+jets background is normalised in this region in the fit. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

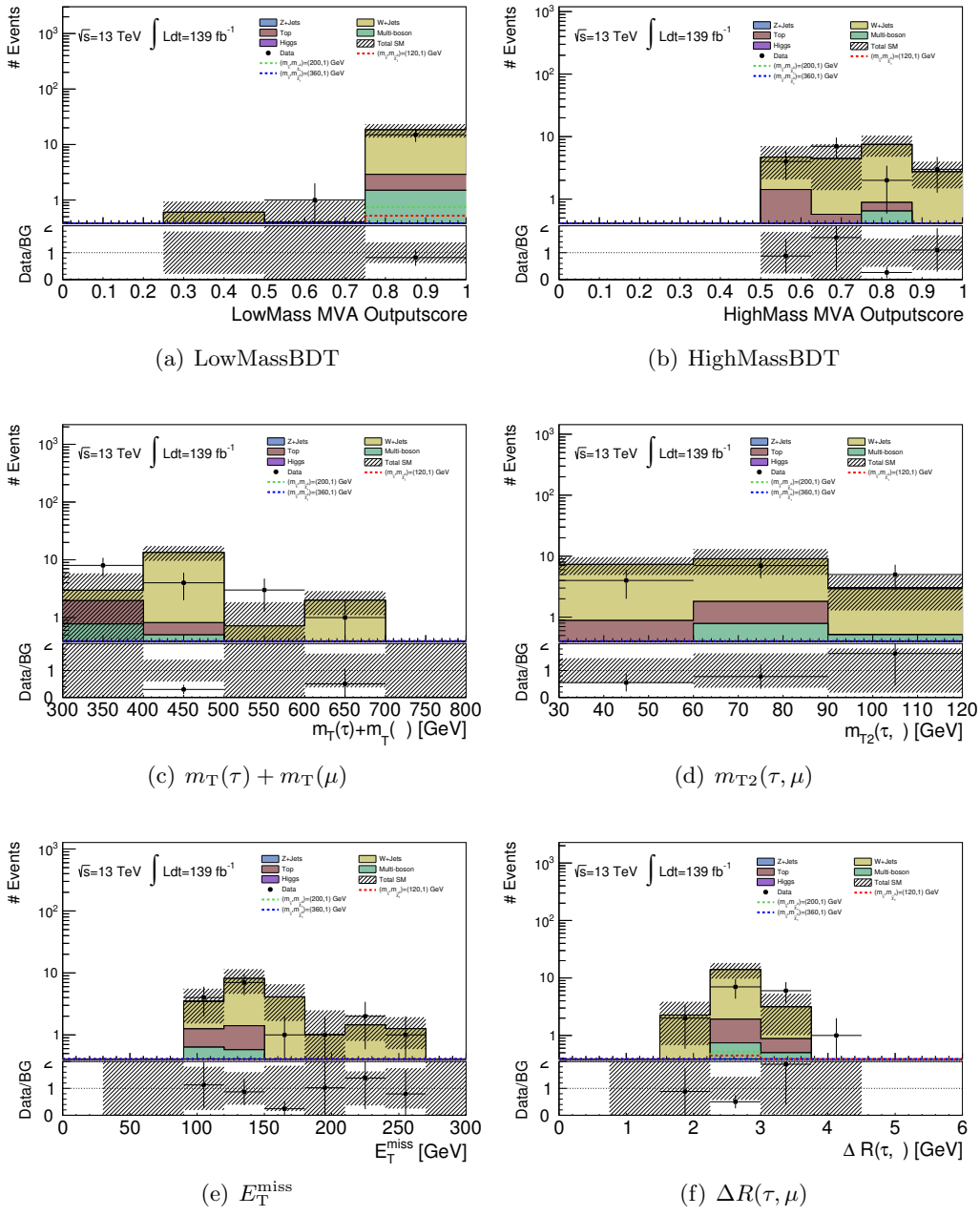


(a) LowMassBDT (b) HighMassBDT

(c) $m_{\mathrm{T}}(\tau) + m_{\mathrm{T}}(\mu)$ (d) $m_{\mathrm{T2}}(\tau, \mu)$

(e) $E_{\mathrm{T}}^{\mathrm{miss}}$ (f) $\Delta R(\tau, \mu)$

Figure E.2.: Pre-fit kinematic distributions for W-VR lowMass. For W-CR, W-VR lowMass and W-VR highMass events containing one signal τ-lepton and one isolated muon are selected. The normalisation for W+jets derived in the fit is applied in this region as a modelling check. This region is however not included in the simultaneous fit itself. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.



(a) LowMassBDT

(b) HighMassBDT

(c) $m_{\mathrm{T}}(\tau) + m_{\mathrm{T}}(\mu)$

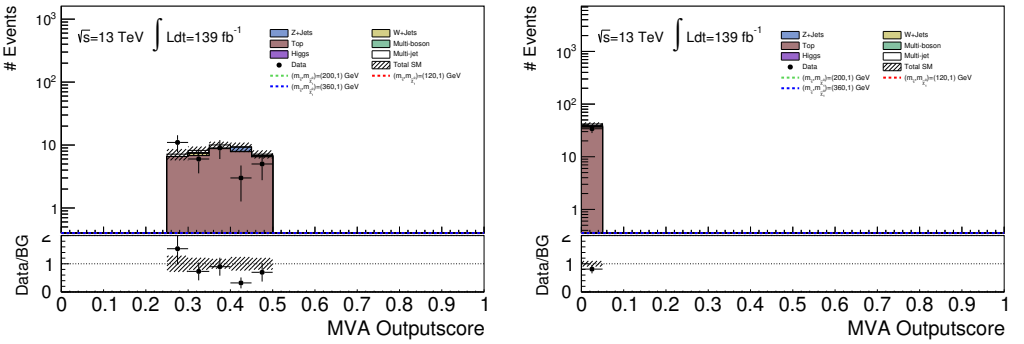
(d) $m_{\mathrm{T2}}(\tau, \mu)$

(e) $E_{\mathrm{T}}^{\mathrm{miss}}$

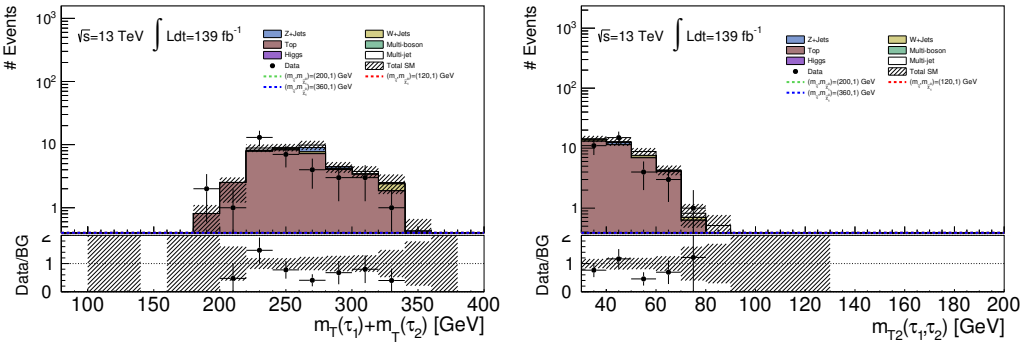
(f) $\Delta R(\tau, \mu)$

Figure E.3.: Pre-fit kinematic distributions for W-VR highMass. For W-CR, W-VR lowMass and W-VR highMass events containing one signal τ-lepton and one isolated muon are selected. The normalisation for W+jets derived in the fit is applied in this region as a modelling check. This region is however not included in the simultaneous fit itself. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation. In subfigure (b) two bins seem to be emptied of W+jets events, which is due to the presence of W+jets events with negative generator weights in these bins.

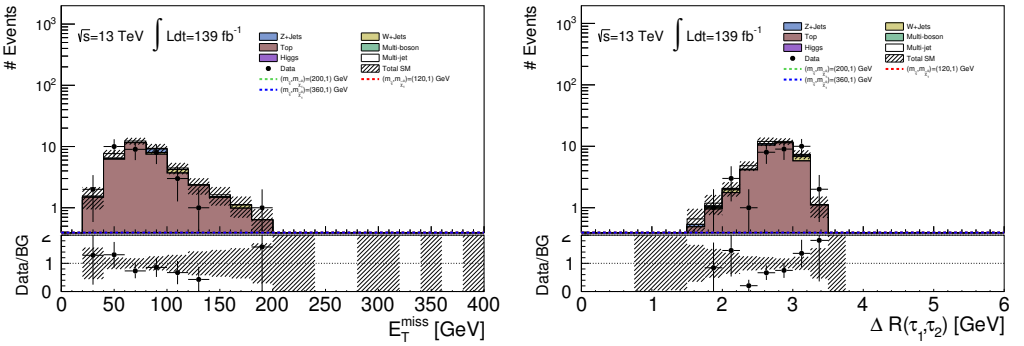
E.2. Top Estimation



(a) LowMassBDT (b) HighMassBDT

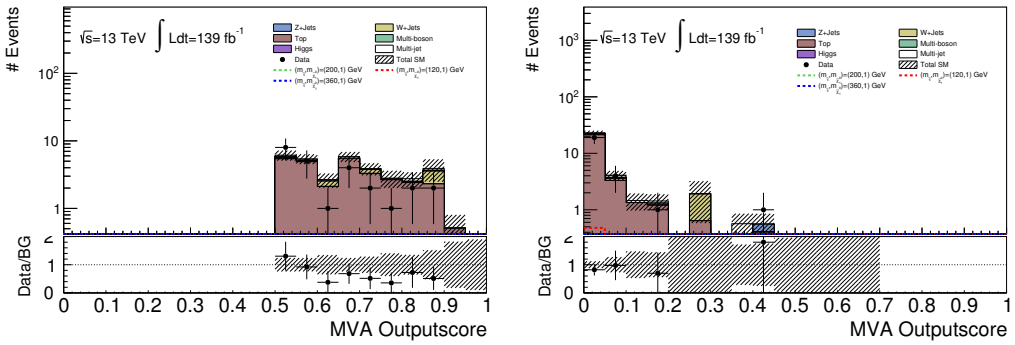


(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$ (d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



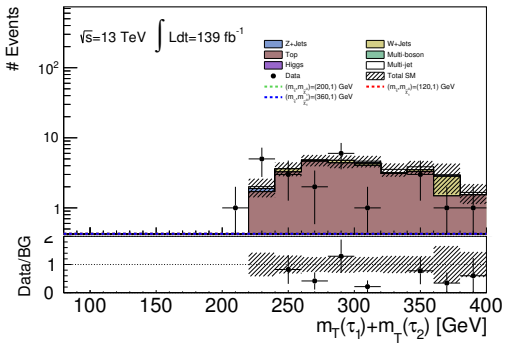
(e) $E_{\mathrm{T}}^{\mathrm{miss}}$ (f) $\Delta R(\tau_1, \tau_2)$

Figure E.4.: Pre-fit kinematic distributions for Top-CR. The Top background is normalised in this region in the fit. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

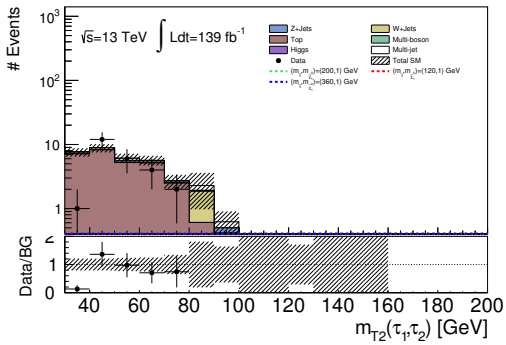


(a) LowMassBDT

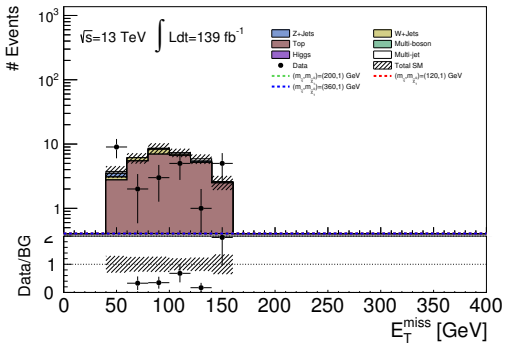
(b) HighMassBDT



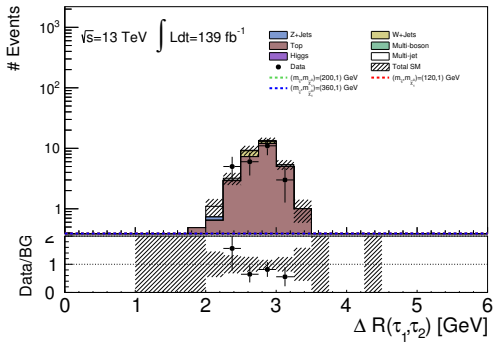
(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$

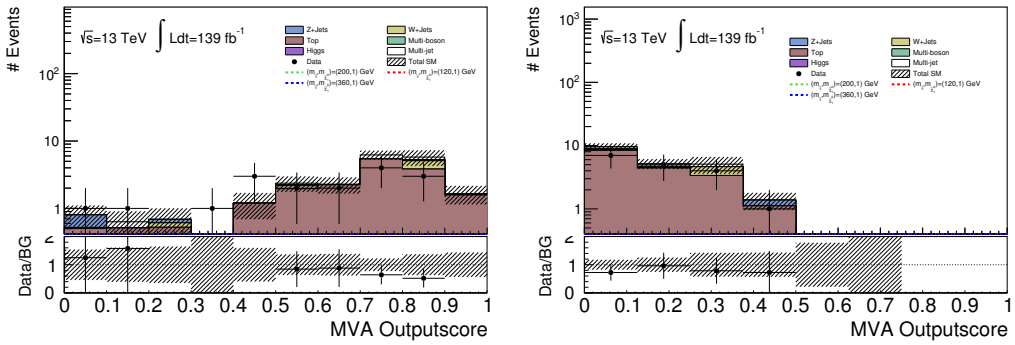


(e) $E_{\mathrm{T}}^{\mathrm{miss}}$



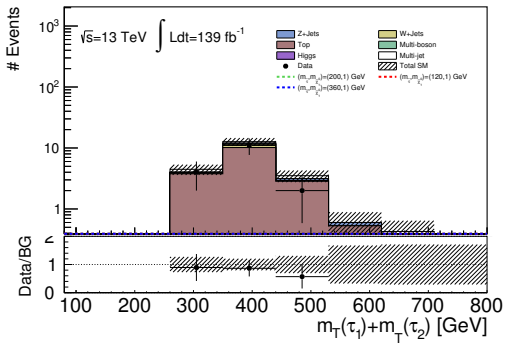
(f) $\Delta R(\tau_1, \tau_2)$

Figure E.5.: Pre-fit kinematic distributions for Top-VR lowMass. The normalisation for Top derived in the fit is applied in this region as a modelling check. This region is however not included in the simultaneous fit itself. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

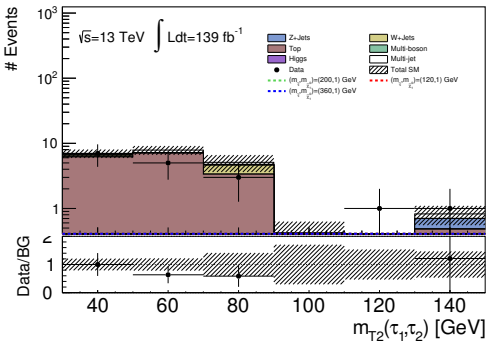


(a) LowMassBDT

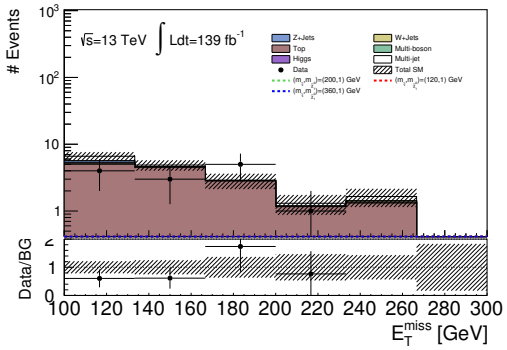
(b) HighMassBDT



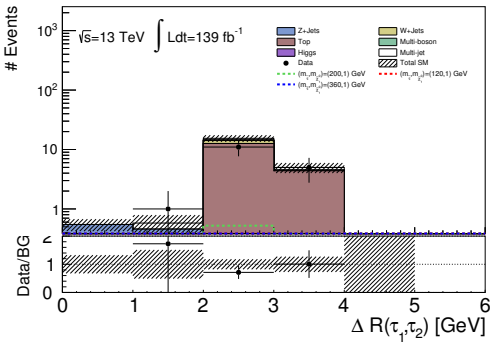
(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



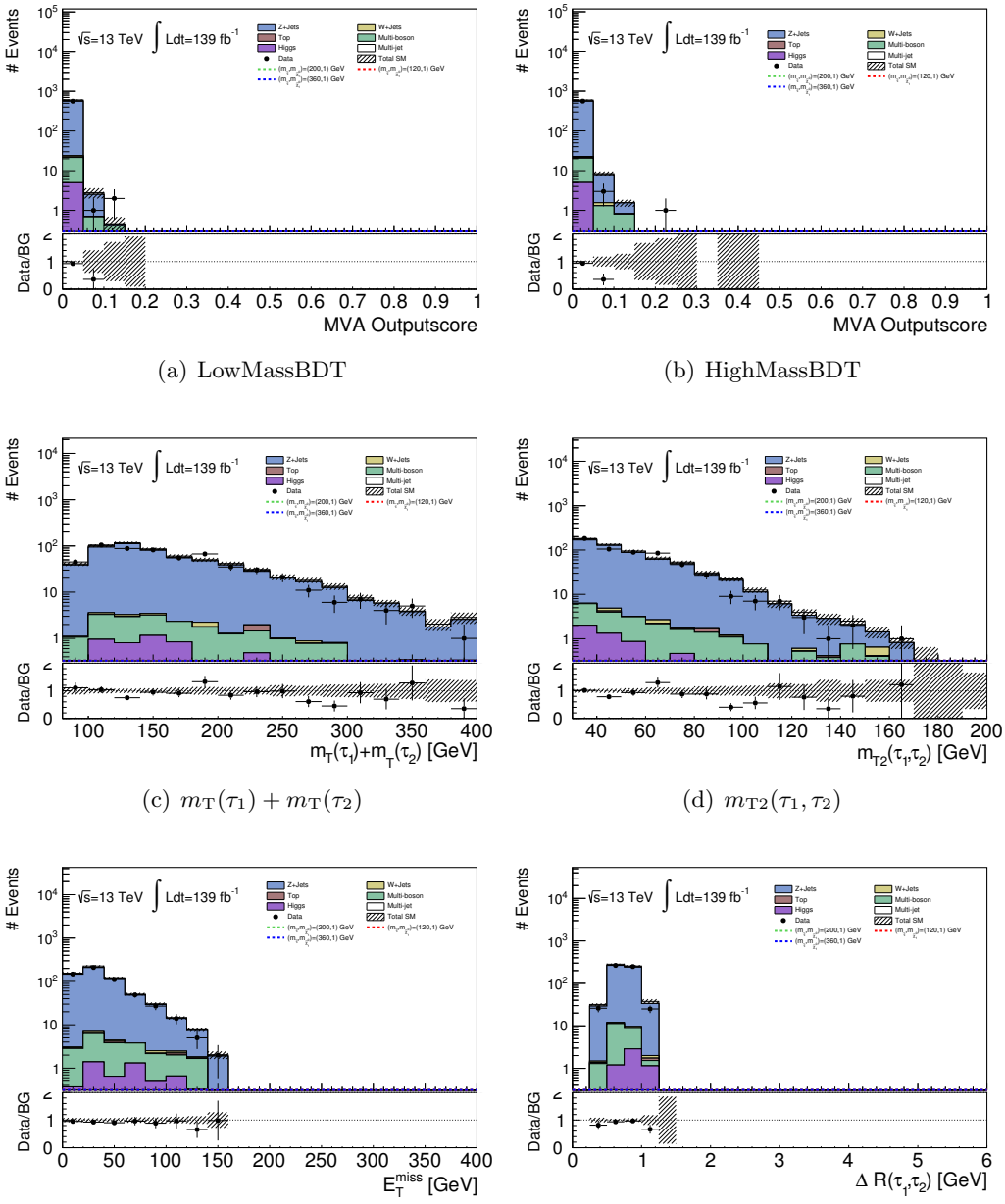
(e) $E_{\mathrm{T}}^{\mathrm{miss}}$



(f) $\Delta R(\tau_1, \tau_2)$

Figure E.6.: Pre-fit kinematic distributions for Top-VR highMass. The normalisation for Top derived in the fit is applied in this region as a modelling check. This region is however not included in the simultaneous fit itself. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

E.3. Z+jets Estimation



(a) LowMassBDT

(b) HighMassBDT

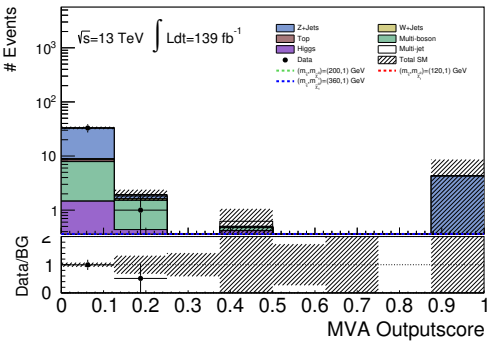
(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$

(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$

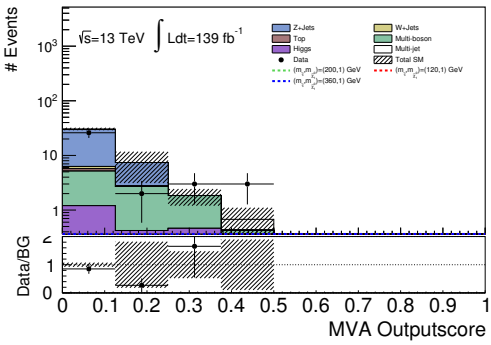
(e) $E_{\mathrm{T}}^{\mathrm{miss}}$

(f) $\Delta R(\tau_1, \tau_2)$

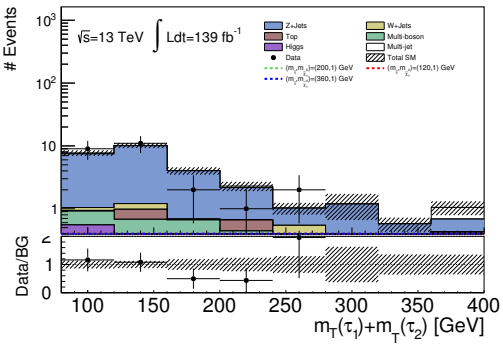
Figure E.7.: Pre-fit kinematic distributions for Z-VR lowMass. This region's purpose is to check the modelling of the Z+jets background close to the lowMass signal regions. The distributions are very well modelled within the statistical uncertainties. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.



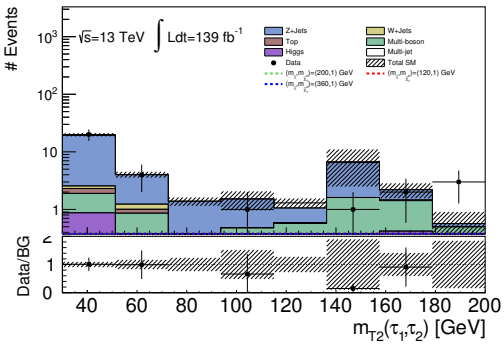
(a) LowMassBDT



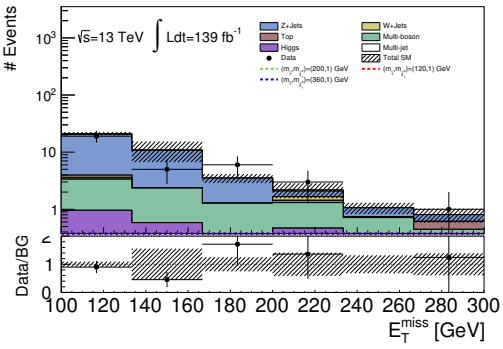
(b) HighMassBDT



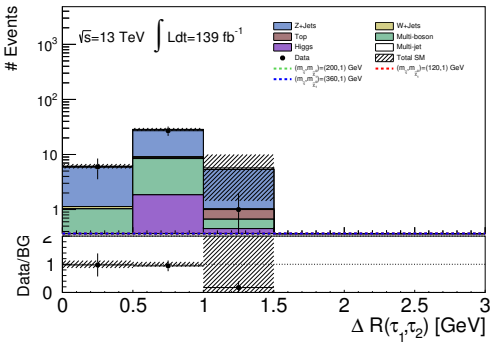
(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



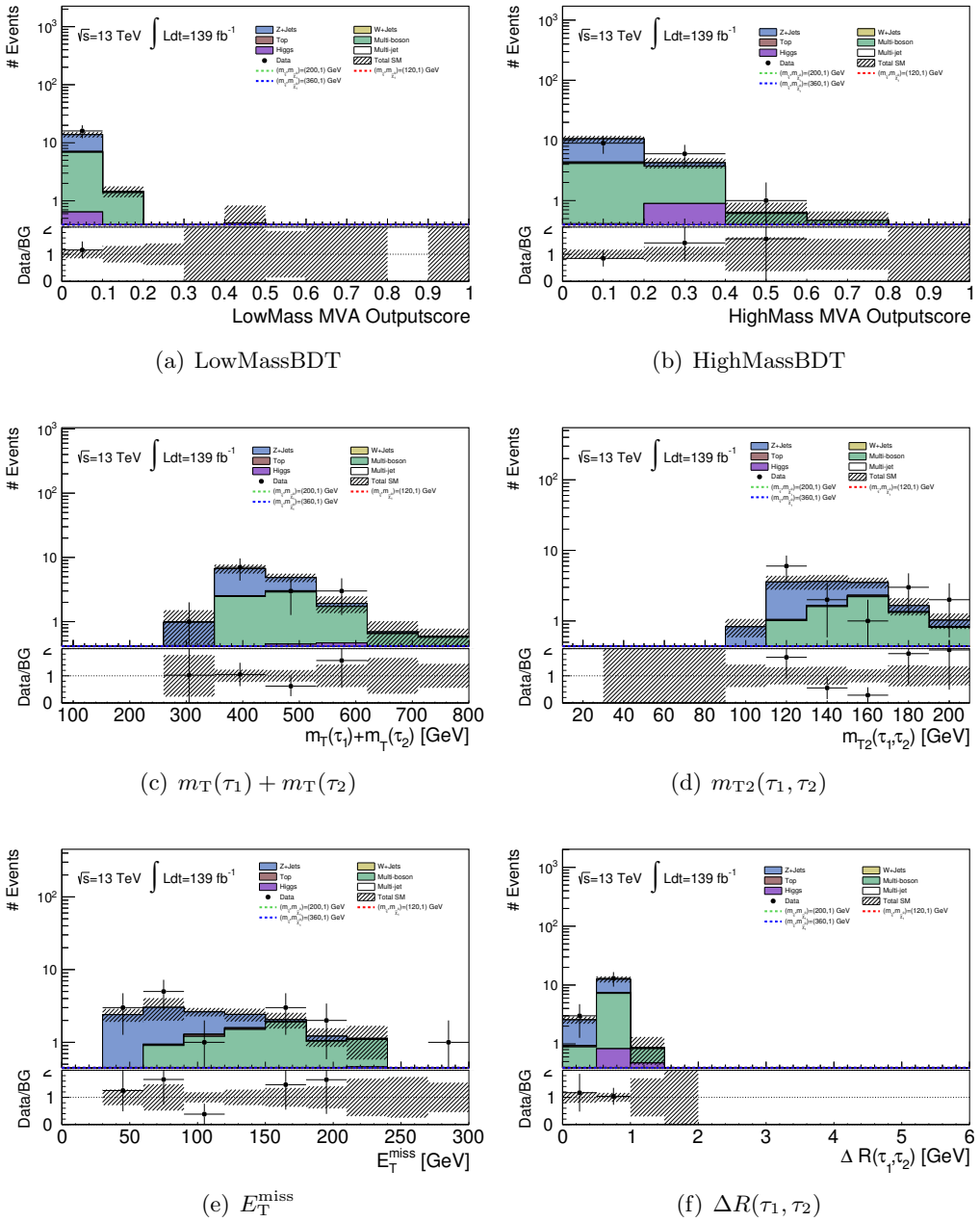
(e) $E_{\mathrm{T}}^{\mathrm{miss}}$



(f) $\Delta R(\tau_1, \tau_2)$

Figure E.8.: Pre-fit kinematic distributions for Z-VR highMass. This region's purpose is to check the modelling of the Z+jets background close to the highMass signal regions. The data is modelled sufficiently well by the Standard Model expectation within the statistical uncertainties. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

E.4. Multiboson Estimation



(a) LowMassBDT (b) HighMassBDT

(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$ (d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$

(e) $E_{\mathrm{T}}^{\mathrm{miss}}$ (f) $\Delta R(\tau_1, \tau_2)$

Figure E.9.: Pre-fit kinematic distributions for the multiboson VR. The multiboson background is the dominant background in SR-highMass. To test its modelling, a validation region is designed as close as possible to SR-highMass while being sufficiently pure in multiboson events. Data and simulated events agree within the statistical uncertainties. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation. The selection provides a relatively pure (but with rather low statistics) multiboson content in this region with some contribution of Z+jets. In principle this could be removed as well by placing a cut on $m_{\mathrm{T2}}(\tau_1, \tau_2)$ or $E_{\mathrm{T}}^{\mathrm{miss}}$. This, however, would even further reduce the statistics in this region. Furthermore, Z+jets is already proven to be sufficiently well modelled in the Z+jets validation regions.

E.5. Multijet Estimation

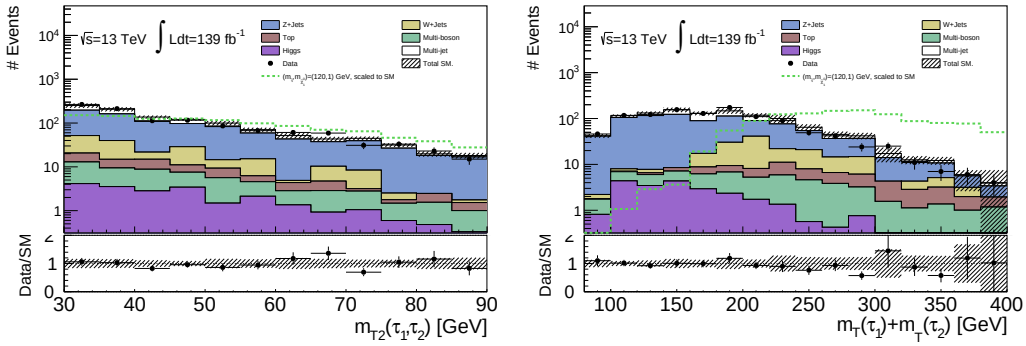
E.5.1. ABCD Method

	CR-A	VR-E	CR-B
Z+jets	533 ± 34	345 ± 56	491 ± 74
W+jets	406 ± 41	236 ± 58	331 ± 55
Top	229.5 ± 6.1	96.8 ± 4.0	94.3 ± 3.9
Multiboson	45.0 ± 2.1	23.5 ± 1.9	22.6 ± 1.5
Higgs	21.4 ± 1.5	7.09 ± 0.46	11.84 ± 0.98
Total MC	1235 ± 54	709 ± 81	951 ± 92
Data	5431	4738	9218
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)$ GeV	22.03 ± 0.50	2.37 ± 0.16	2.10 ± 0.15
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)$ GeV	41.5 ± 1.3	11.77 ± 0.70	12.05 ± 0.72
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)$ GeV	4.36 ± 0.25	0.309 ± 0.066	0.172 ± 0.048

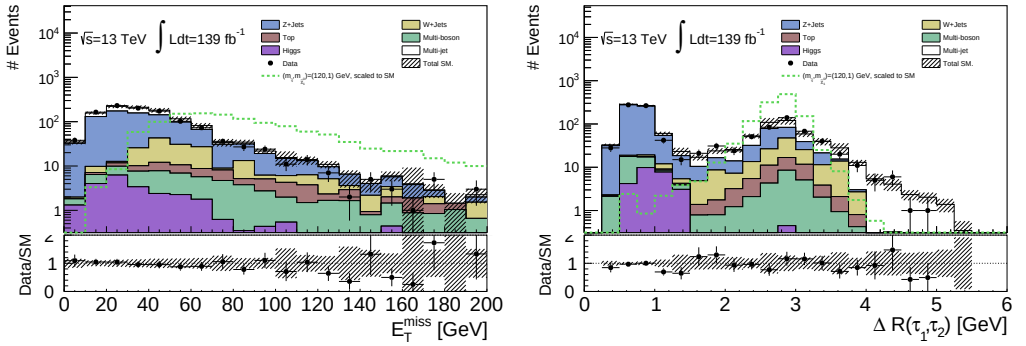
	pre-SR-D	VR-F	CR-C
Z+jets	1552 ± 51	1055 ± 73	1840 ± 150
W+jets	273 ± 36	214 ± 54	266 ± 61
Top	399.8 ± 8.1	171.3 ± 5.3	187.8 ± 5.6
Multiboson	110.0 ± 3.1	46.5 ± 2.2	57.0 ± 2.4
Higgs	57.0 ± 2.6	22.8 ± 1.5	39.6 ± 2.4
Total MC	2391 ± 63	1509 ± 91	2390 ± 160
Data	3107	2006	3708
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (200, 1)$ GeV	88.65 ± 0.99	11.07 ± 0.36	10.43 ± 0.34
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)$ GeV	163.2 ± 2.6	45.4 ± 1.4	46.4 ± 1.4
$(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)$ GeV	19.12 ± 0.51	1.09 ± 0.12	0.96 ± 0.12

Table E.1.: Monte Carlo and data event yields in the regions used for the multijet estimation with the ABCD method. The uncertainties correspond to the statistical uncertainty only.

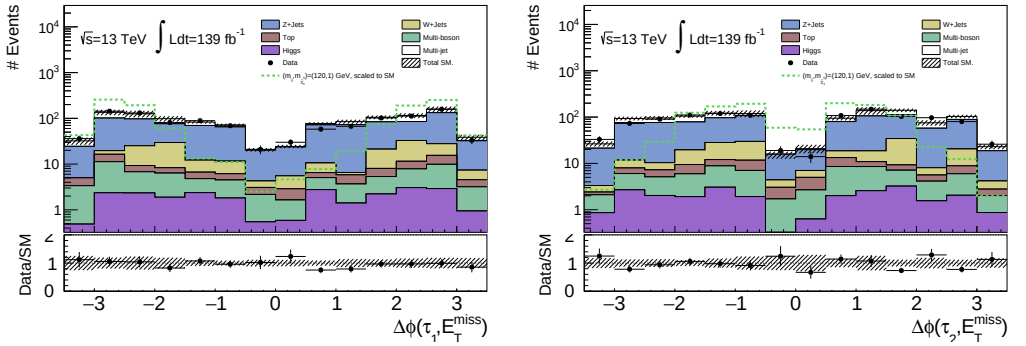
E.5.2. Pre-Signal Region SR-D with Unsimplified ABCD Method



(a) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ (b) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



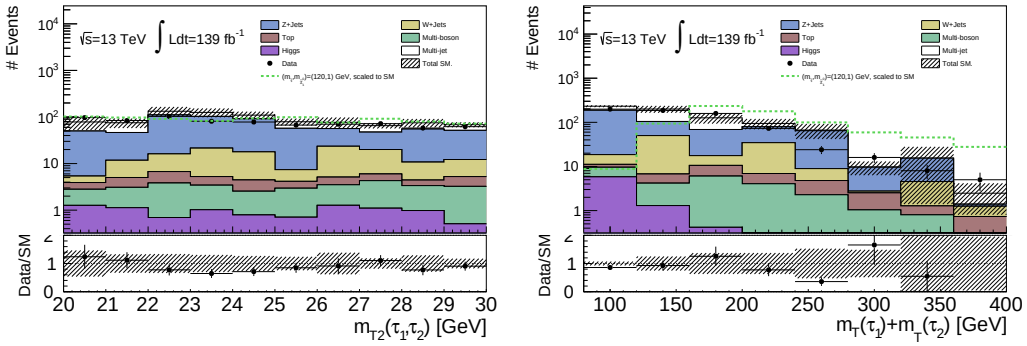
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$ (d) $\Delta R(\tau_1, \tau_2)$



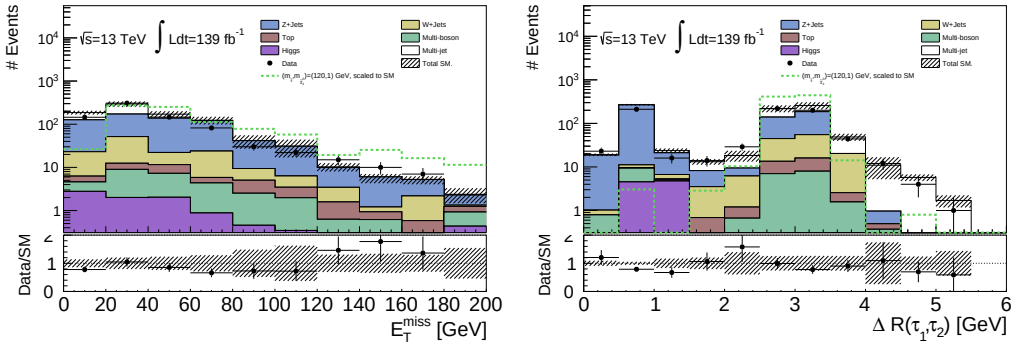
(e) $\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure E.10.: Pre-fit distributions in the pre-SR-D with unsimplified multijet estimate. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

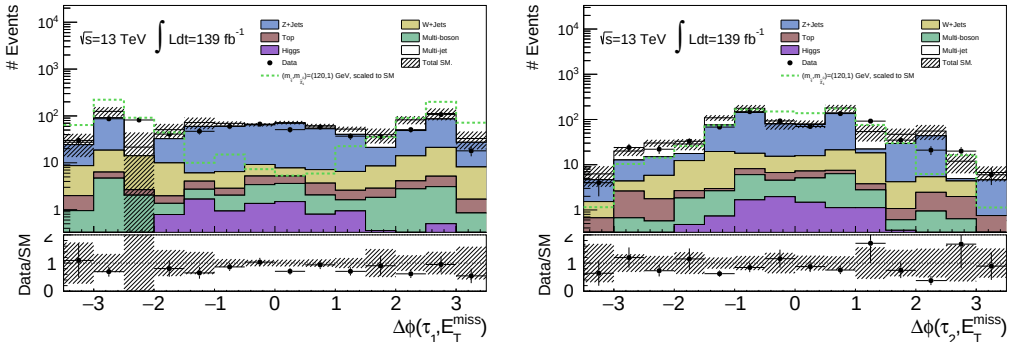
E.5.3. Validation Region VR-F with Unsimplified ABCD Method



(a) $m_{\mathrm{T}2}(\tau_1, \tau_2)$ (b) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



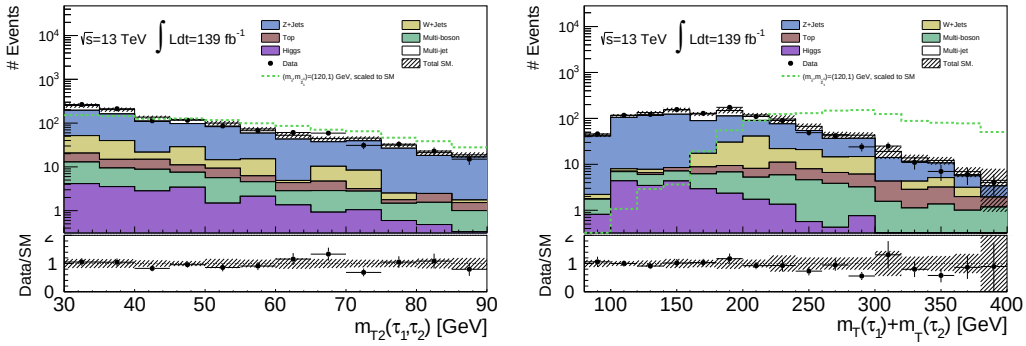
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$ (d) $\Delta R(\tau_1, \tau_2)$



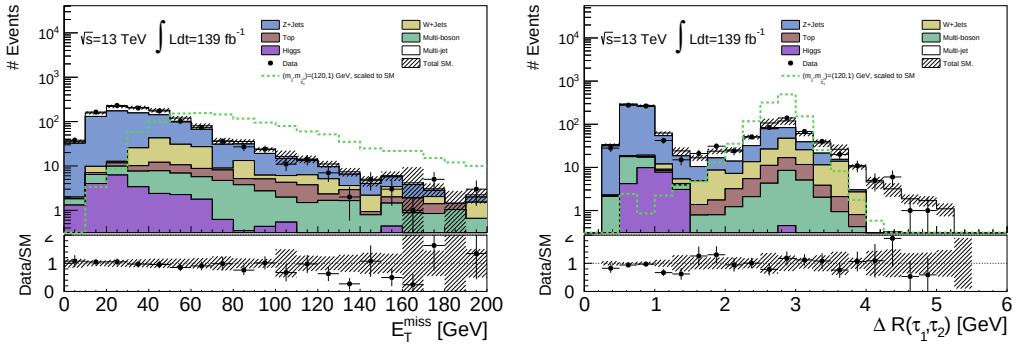
(e) $\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure E.11.: Pre-fit distributions in VR-F with unsimplified multijet estimate. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

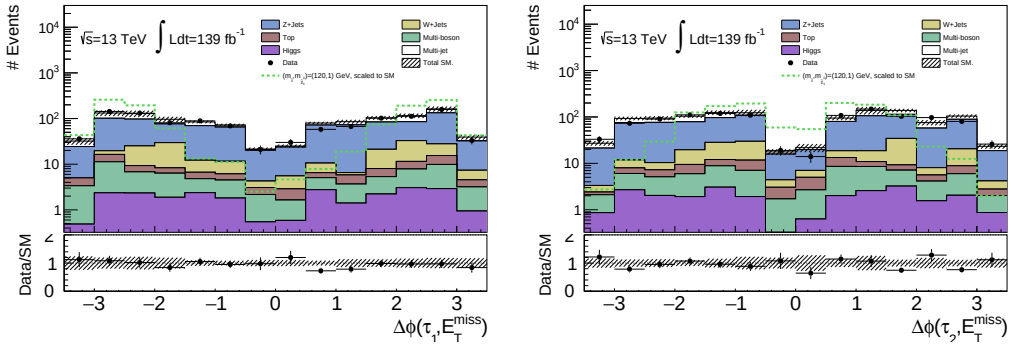
E.5.4. Pre-Signal Region SR-D with Simplified ABCD Method



(a) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ (b) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



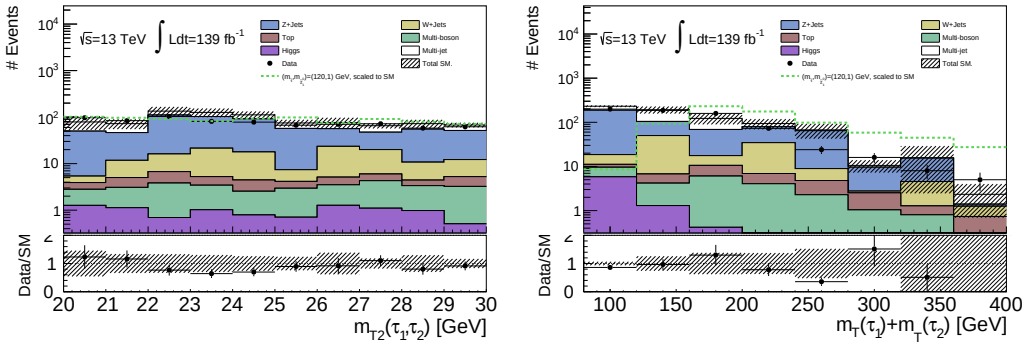
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$ (d) $\Delta R(\tau_1, \tau_2)$



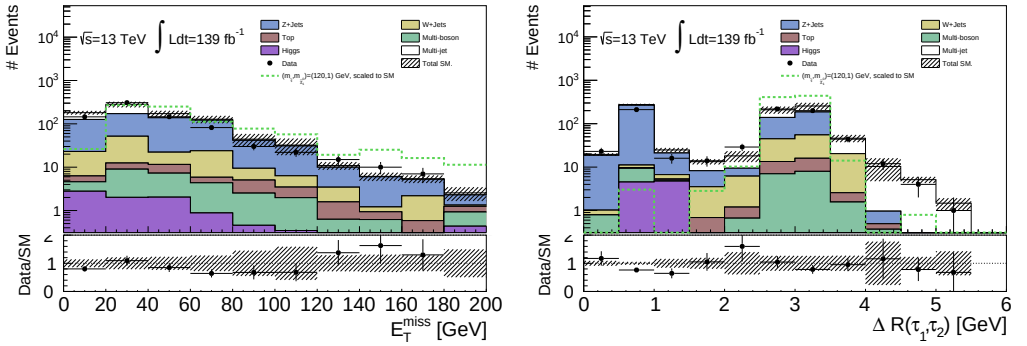
(e) $\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure E.12.: Pre-fit distributions in the pre-SR-D with simplified multijet estimate. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

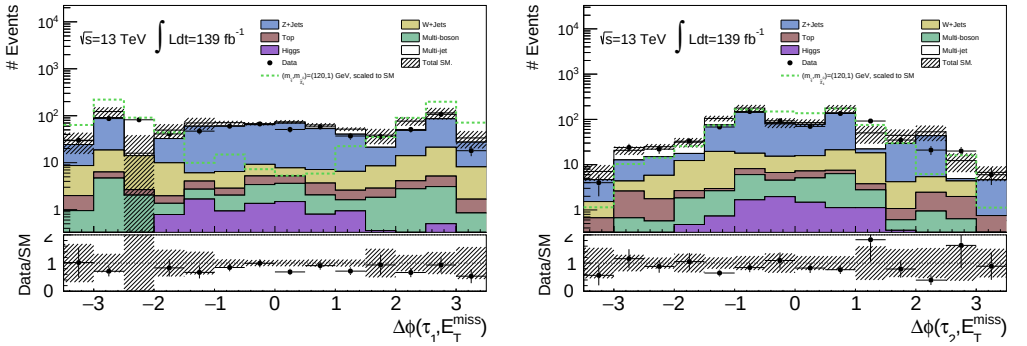
E.5.5. Validation Region VR-F with Simplified ABCD Method



(a) $m_{\mathrm{T2}}(\tau_1, \tau_2)$ (b) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$



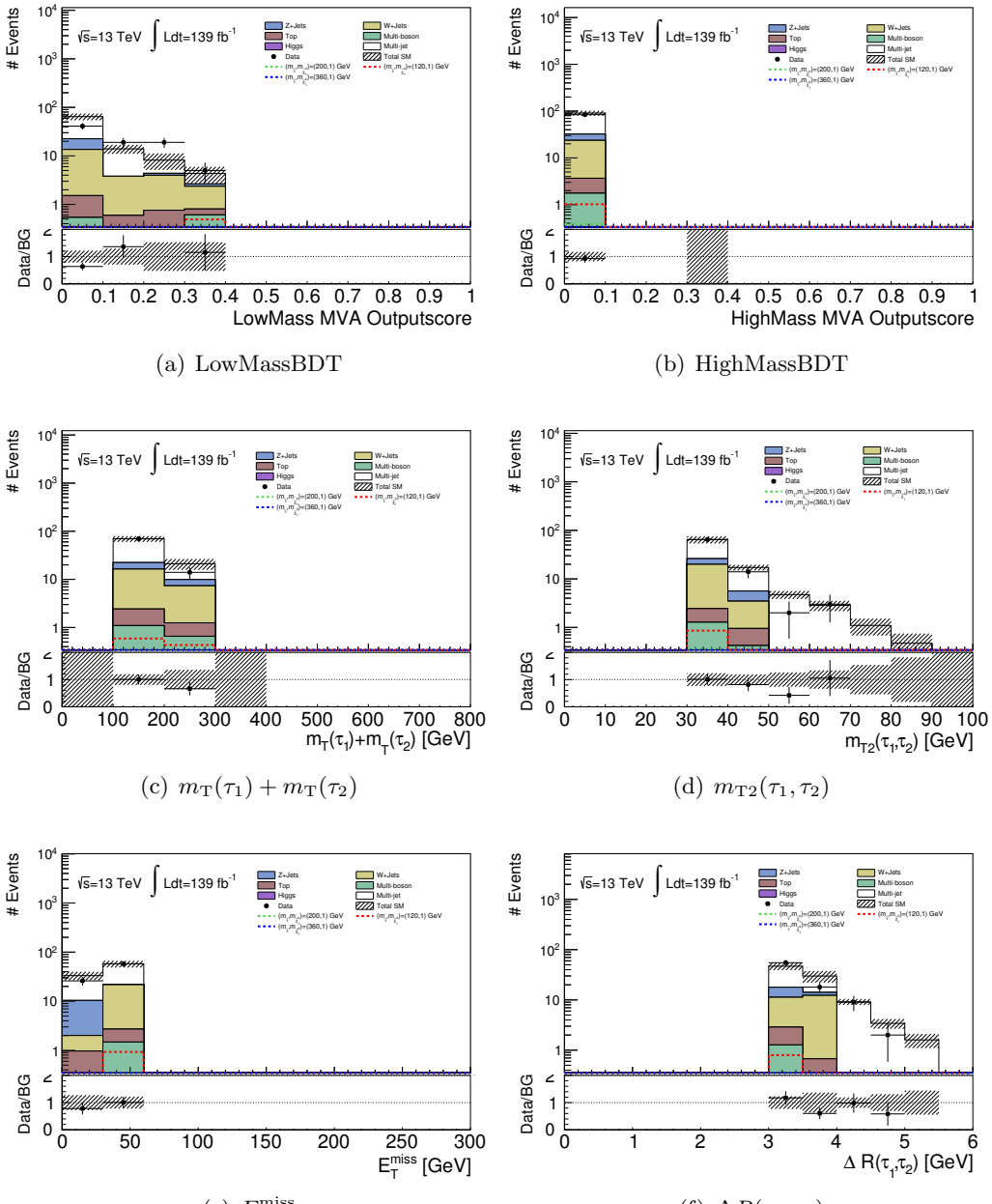
(c) $E_{\mathrm{T}}^{\mathrm{miss}}$ (d) $\Delta R(\tau_1, \tau_2)$



(e) $\Delta\phi(\tau_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $\Delta\phi(\tau_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure E.13.: Pre-fit distributions in VR-F with simplified multijet estimate. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

E.5.6. Multijet Validation MJVR



(a) LowMassBDT

(b) HighMassBDT

(c) $m_{\mathrm{T}}(\tau_1) + m_{\mathrm{T}}(\tau_2)$

(d) $m_{\mathrm{T2}}(\tau_1, \tau_2)$

(e) $E_{\mathrm{T}}^{\mathrm{miss}}$

(f) $\Delta R(\tau_1, \tau_2)$

Figure E.14.: Pre-fit kinematic distributions for the multijet-VR. The multijet background is one of the most dominant backgrounds in SR-lowMass. Its modelling in close proximity to SR-lowMass is tested in this region, which is not only very pure in multijet events but also shows good agreement of data and expected Standard Model events. No additional cut on the highMass BDT is applied. The shaded areas correspond to the statistical uncertainties on the total Standard Model expectation.

F. Second Search: Additional Material on Fits

F.1. Background-Only Fit

In the following section, more detailed information on the background-only fit results is given. Fig. F.2 shows the behaviour of the nuisance parameters during the fit. The black points correspond to the central values of the fit parameters for the individual sources of systematic uncertainties ("α") while the blue points show the normalization factors for W+jets and Top ("μ") as well as nuisance parameters to account for the statistical uncertainties in the control regions ("γ"). No "pull" of individual systematic uncertainties is observed as the deviations of estimated and observed yields are predominantly covered by the normalization factors for the W+jets and Top background categories. Fig. F.1 gives information on the correlations between the most important nuisance parameters that are encountered in the fit. The correlations are mostly small. The normalization factors are however anti-correlated with the available statistics in the respective control region, which is to be expected. The nuisance parameters listed in the correlation matrix will be explained in more detail in the following:

- alpha_JET_JER_EffectiveNP_1: Uncertainty on jet energy resolution.
- alpha_JET_Pileup_Offset_Mu: Pileup jet related uncertainty.
- alpha_JET_PunchThrough_MC16: Uncertainty taking into account punch-through of jets into the muon detector system.
- alpha_MET_SoftTrk_ResoPara: Uncertainty on resolution of the soft term in the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$.
- alpha_ZjetsTheoSys: Theoretical uncertainty of the Z+jets background, using the estimate of the first iteration of the stau search for each region.
- gamma_stat_TCR_cuts_bin_0: Statistical uncertainty of the background estimate in the Top control region.
- gamma_stat_WCR_cuts_bin_0: Statistical uncertainty of the background estimate in the W+jets control region.
- mu_Top: Normalization factor of the Top background, derived in the Top control region.
- mu_W: Normalization factor of the W+jets background, derived in the W+jets control region.

In the tables Tab. F.1 to Tab. F.3 the number of events for each background before and after the fit is listed for each region together with the observed number of events.

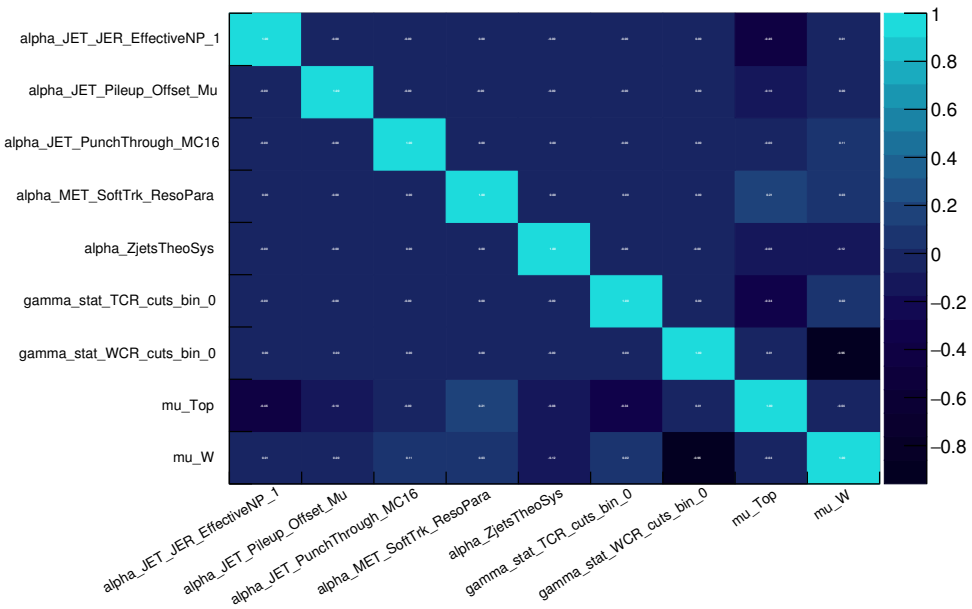


Figure F.1.: Reduced correlation matrix of selected nuisance parameters in the background only fit. Most of the correlations between the individual nuisance parameters are small. An exception is the observed anti-correlation of the normalization factors for Top and W+jets and the statistical uncertainties in the respective control region.

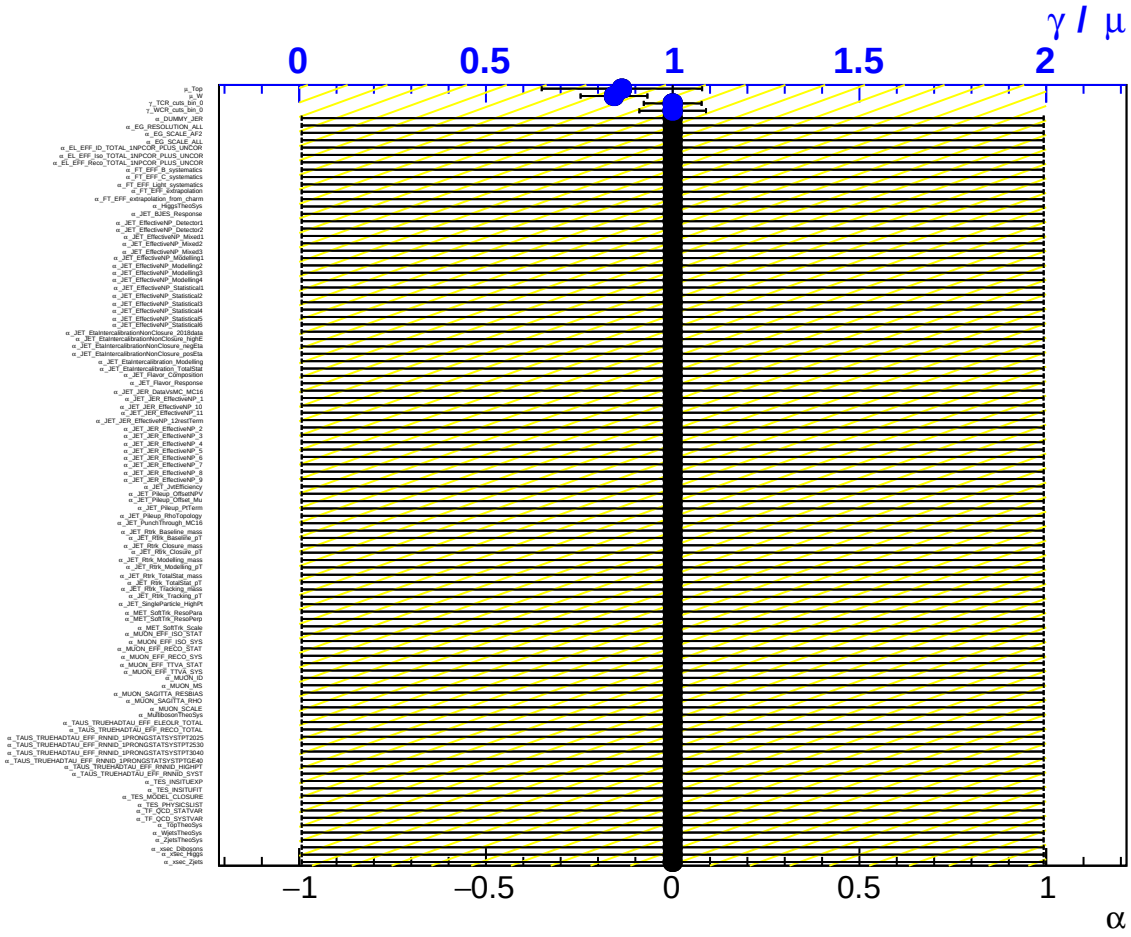


Figure F.2.: Nuisance parameter pulls for the background-only fit. The nuisance parameters indicated by the black points all correspond to systematic uncertainties as outlined in Sec. 6.4 and Sec. 7.5.1. The upper two blue points stand for the derived normalization factors for the W+jets and Top backgrounds. The other blue points correspond to nuisance parameters introduced to account for the statistical uncertainties in the two control regions. No pulls are observed for any of the nuisance parameters other than the normalization factors as deviations between Standard Model expectation and data are covered predominantly by those two parameters.

	TCR	TVR-lowMass	TVR-highMass
Observed events	34	25	17
Post-fit background events	34.0 ± 5.9	28.1 ± 7.2	17.8 ± 4.3
Post-fit Z+jets events	$1.6^{+2.7}_{-1.6}$	$0.29^{+0.97}_{-0.29}$	0.76 ± 0.29
Post-fit W+jets events	$0.7^{+2.6}_{-0.7}$	2.24 ± 0.70	1.39 ± 0.28
Post-fit Top events	31.3 ± 7.3	25.3 ± 7.0	14.9 ± 4.2
Post-fit Multiboson events	0.230 ± 0.079	0.150 ± 0.055	0.36 ± 0.10
Post-fit Higgs events	0.058 ± 0.031	0.041 ± 0.016	0.17 ± 0.14
Post-fit Multijet	0.159 ± 0.043	–	0.226 ± 0.063
Pre-fit Standard Model events	39.1 ± 6.0	32.5 ± 4.2	20.4 ± 2.7
Pre-fit Z+jets events	$1.6^{+2.7}_{-1.6}$	$0.29^{+0.98}_{-0.29}$	0.76 ± 0.29
Pre-fit W+jets events	$0.8^{+3.1}_{-0.8}$	2.66 ± 0.78	1.65 ± 0.27
Pre-fit Top events	36.2 ± 2.8	29.3 ± 4.0	17.2 ± 2.4
Pre-fit Multiboson events	0.230 ± 0.080	0.150 ± 0.055	0.36 ± 0.10
Pre-fit Higgs events	0.058 ± 0.031	0.041 ± 0.016	0.17 ± 0.14
Pre-fit Multijet	0.159 ± 0.043	–	0.226 ± 0.064

Table F.1.: Events yields pre- and post-fit for Top CR and VRs. The given uncertainties correspond to the combined statistical and systematic uncertainties.

	WCR	WVR-lowMass	WVR-highMass
Observed events	4138	357	16
Post-fit background events	4138 ± 64	330 ± 69	16.6 ± 5.4
Post-fit Z+jets events	249 ± 58	5.4 ± 2.5	0.061 ± 0.048
Post-fit W+jets events	3650 ± 120	284 ± 68	13.6 ± 4.9
Post-fit Top events	91 ± 28	17.3 ± 5.8	1.38 ± 0.52
Post-fit Multiboson events	145 ± 55	22.8 ± 4.7	1.61 ± 0.54
Post-fit Higgs events	3.79 ± 0.87	$0.8^{+1.3}_{-0.8}$	0.0016 ± 0.0011
Post-fit Multijet events	–	–	–
Pre-fit Standard Model events	4830 ± 440	386 ± 73	19.35 ± 6.2
Pre-fit Z+jets events	249 ± 59	5.4 ± 2.5	0.061 ± 0.049
Pre-fit W+jets events	4330 ± 390	337 ± 71	16.1 ± 5.6
Pre-fit Top events	105 ± 19	20.0 ± 4.5	1.59 ± 0.46
Pre-fit Multiboson events	145 ± 55	22.8 ± 4.7	1.61 ± 0.54
Pre-fit Higgs events	3.79 ± 0.87	$0.8^{+1.3}_{-0.8}$	0.0016 ± 0.0011
Pre-fit Multijet events	–	–	–

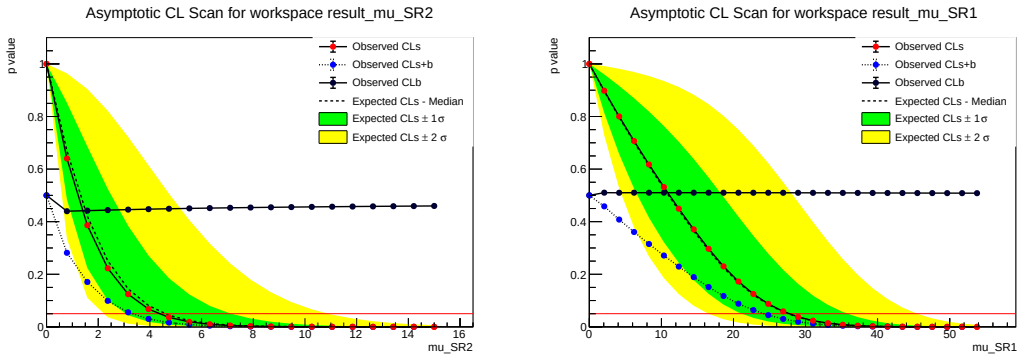
Table F.2.: Events yields pre- and post-fit for W+jets CR and VRs. The given uncertainties correspond to the combined statistical and systematic uncertainties.

	ZVR-lowMass	ZVR-highMass	MBVR	MJVR
Observed events	565	34	16	84
Post-fit background events	590 ± 110	39 ± 14	15.9 ± 4.7	87 ± 26
Post-fit Z+jets events	550 ± 100	28 ± 14	6.8 ± 3.7	8^{+14}_{-8}
Post-fit W+jets events	1.54 ± 0.67	0.41 ± 0.12	–	16.9 ± 7.1
Post-fit Top events	0.57 ± 0.24	0.52 ± 0.17	0.216 ± 0.065	1.66 ± 0.78
Post-fit Multiboson events	18 ± 17	7.8 ± 2.2	7.6 ± 1.6	1.75 ± 0.44
Post-fit Higgs events	5.2 ± 1.2	2.32 ± 0.73	1.31 ± 0.61	$0.0082^{+0.0096}_{-0.082}$
Post-fit Multijet events	14.2 ± 3.7	–	–	59 ± 17
Pre-fit Standard Model events	590 ± 110	39 ± 15	16.0 ± 4.7	91 ± 27
Pre-fit Z+jets events	550 ± 100	28 ± 14	6.8 ± 3.7	9^{+14}_{-9}
Pre-fit W+jets events	1.82 ± 0.77	0.49 ± 0.14	–	20.0 ± 8.3
Pre-fit Top events	0.66 ± 0.22	0.60 ± 0.13	0.250 ± 0.043	1.92 ± 0.76
Pre-fit Multiboson events	18^{+18}_{-18}	7.8 ± 2.2	7.6 ± 1.6	1.75 ± 0.44
Pre-fit Higgs events	5.2 ± 1.2	2.32 ± 0.73	1.31 ± 0.62	$0.0082^{+0.0096}_{-0.0082}$
Pre-fit Multijet events	14.2 ± 3.7	–	–	59 ± 17

Table F.3.: Events yields pre- and post-fit for the Z+jets, Multiboson and Multijet validation regions. The given uncertainties correspond to the combined statistical and systematic uncertainties.

F.2. Model Independent Upper Limit Scans

In Fig. F.3 the scans over various values for the signal strength μ_{SR} are shown for SR-highMass and the discovery SR-lowMass regions. The signal strength corresponds in this case to the number of additional events not predicted by the background estimate in one of the signal regions. The signal strength at which the CL_s value drops below 0.05 marks the number of additional events that can be excluded at 95%, i.e. the signal+background hypothesis can be rejected. These are still expected limits as the observed data is not yet viewed.



(a) SR-highMass (b) SR-lowMass (discovery)

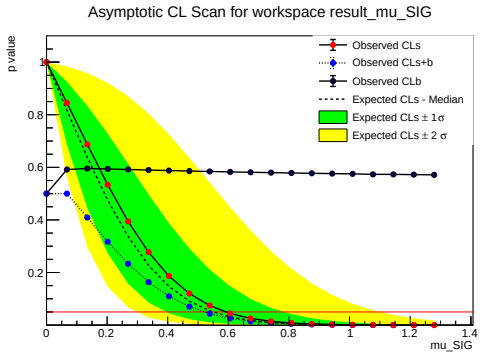
Figure F.3.: Model independent upper limit scans. The value of μ_{SR} at which the CL_s-value becomes smaller than 0.05 corresponds to the number of additional events in the respective signal region that can be excluded at 95%.

F.3. Model Dependent Upper Limit Scans

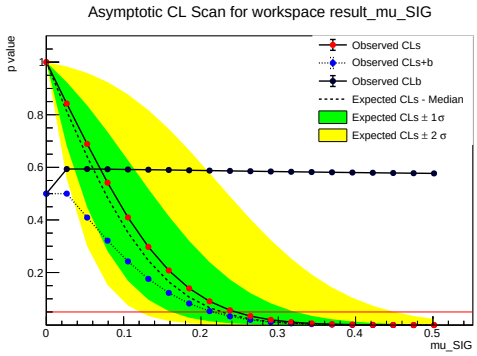
Fig. F.4 shows the outcome of the upper limit scans for some benchmark signal models with the combination of $\tilde{\tau}_L$ and $\tilde{\tau}_R$ pair production. The red points show the development of the CL_s value, where it should be noted that only the background estimation has been used for its derivation and no observed data. The green (yellow) bands correspond to the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainty. The 0.05 threshold for the CL_s value is indicated by a horizontal red line. If the CL_s value passes this threshold with the signal strength being still < 1, the corresponding signal model is considered excluded at 95% CL. From the shown plots it can be seen that this is the case for stau masses of 80 GeV and 120 GeV with near to massless neutralino. For a stau mass of 360 GeV and a lightest neutralino of 160 GeV the model cannot be excluded considering the CL_s value, but the -1σ band (green band) drops below 0.05 at a signal strength < 1. A stau mass of 520 GeV with near to massless neutralino cannot be excluded from its expected CL_s value but it can within $\pm 2\sigma$.

For the production of $\tilde{\tau}_L$ pairs separately, the outcome of the upper limit scans for a few exemplary models are shown in Fig. F.5. Equivalently, some of the upper limit scans for separate $\tilde{\tau}_R$ production are given in Fig. F.6. For stau masses of 120 GeV and 160 GeV with near to massless neutralino the model can be excluded at 95% even within the $\pm 1\sigma$ uncertainty band.

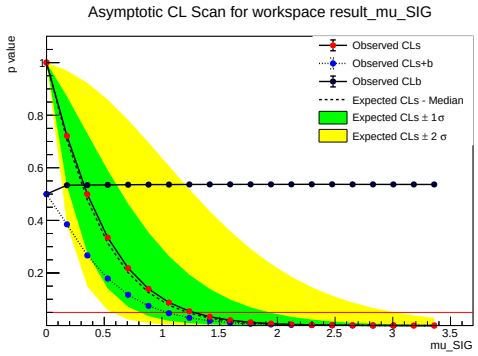
All of the mentioned exclusion limits are, however, still expected limits as the observed data has not yet been compared to the Standard Model expectation in the signal regions.



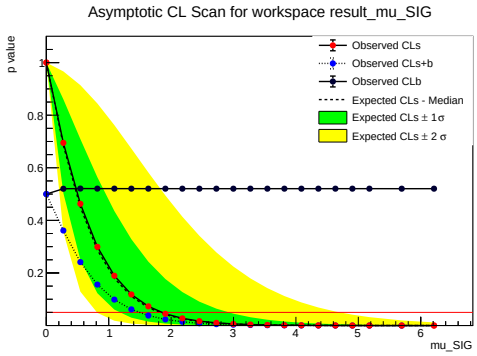
(a) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (80, 1)\,\text{GeV}$



(b) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$

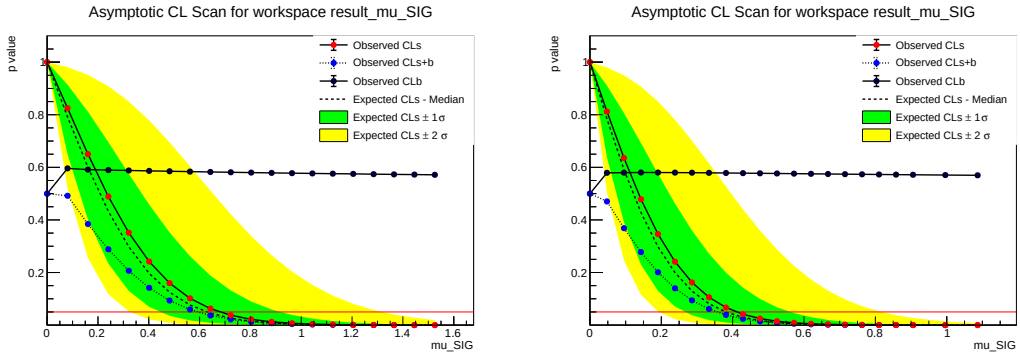


(c) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 160)\,\text{GeV}$



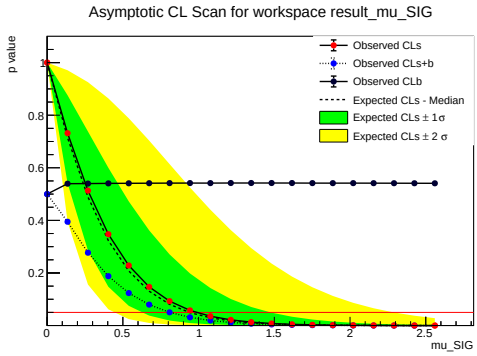
(d) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (520, 1)\,\text{GeV}$

Figure F.4.: Upper limit scans for four benchmark signal models with combined stau eigenstate scenarios. A model is considered excluded at 95% CL if the CL_s value (red dotted line) drops below 0.05 (red horizontal line) at a signal strength less than 1.

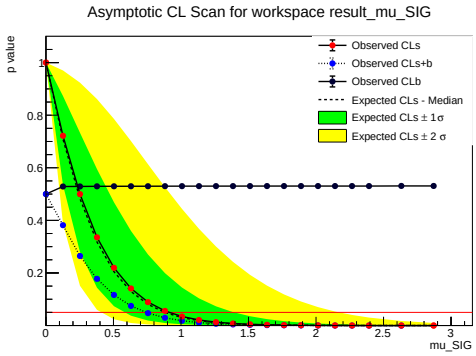


(a) $(m_{\tilde{\tau}_L}, m_{\tilde{\chi}_1^0}) = (100, 1)\,\text{GeV}$

(b) $(m_{\tilde{\tau}_L}, m_{\tilde{\chi}_1^0}) = (160, 1)\,\text{GeV}$

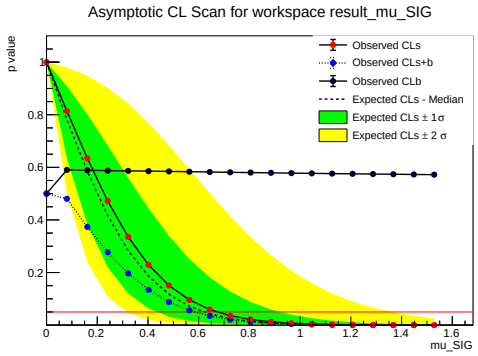


(c) $(m_{\tilde{\tau}_L}, m_{\tilde{\chi}_1^0}) = (280, 80)\,\text{GeV}$

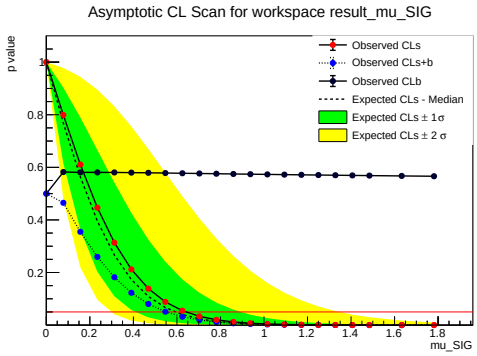


(d) $(m_{\tilde{\tau}_L}, m_{\tilde{\chi}_1^0}) = (340, 1)\,\text{GeV}$

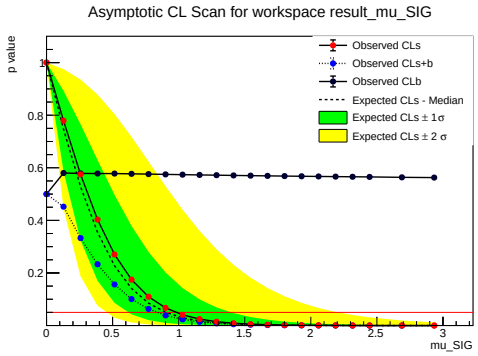
Figure F.5.: Upper limit scans for four benchmark signal models for $\tilde{\tau}_L$ pair production. A model is considered excluded at 95% CL if the CL_s value (red dotted line) drops below 0.05 (red horizontal line) at a signal strength less than 1.



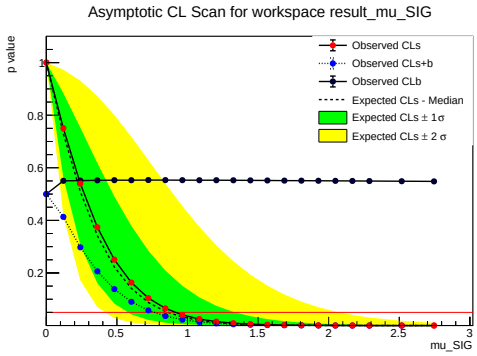
(a) $(m_{\tilde{\tau}_R}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$



(b) $(m_{\tilde{\tau}_R}, m_{\tilde{\chi}_1^0}) = (160, 1)\,\text{GeV}$



(c) $(m_{\tilde{\tau}_R}, m_{\tilde{\chi}_1^0}) = (160, 40)\,\text{GeV}$



(d) $(m_{\tilde{\tau}_R}, m_{\tilde{\chi}_1^0}) = (220, 1)\,\text{GeV}$

Figure F.6.: Upper limit scans for four benchmark signal models for $\tilde{\tau}_R$ pair production. A model is considered excluded at 95% CL if the CL_s value (red dotted line) drops below 0.05 (red horizontal line) at a signal strength less than 1.

G. Simulation Bug Affecting Signal Samples

Shortly before the finalization of this thesis, it became known that the signal simulations of the second iteration of the stau pair production signal were affected by a bug concerning the Pythia merging and the assigned Monte Carlo weights. A correction was implemented such that four benchmark models could be tested with the new version of the ATLAS simulation software (referred to as Release 21.6.46) and compared to the previous version (in the following called Release 21.6.27) at truth level. The chosen benchmark models are $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)$, $(120, 60)$, $(360, 1)$ and $(360, 160)$ GeV. This way, the new version of the ATLAS simulation software containing the bug correction can be tested in the most important kinematic regions of the signal models.

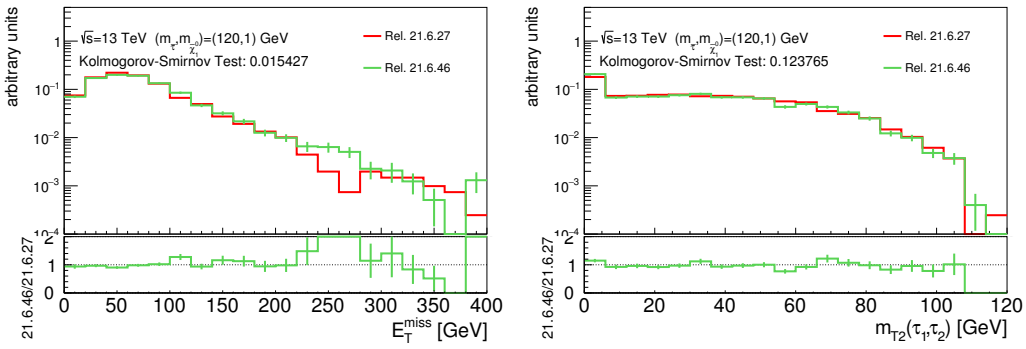
Some of the most important kinematic variables are checked for each of the benchmark models for both software releases. A Kolmogorov-Smirnov test is performed in addition, to get a quantitative estimate of the agreement of the two distributions. The output of the KS-test is automatically converted into a p-value by ROOT [111]. If the p-value is > 0.05, the hypothesis of the two distributions being identical cannot be rejected. Otherwise, if the p-value is very small, the two distributions are incompatible with each other. The plots are shown in Fig. G.1 to Fig. G.4. For high stau masses there is good agreement between the two software releases while there are some discrepancies for the lower stau masses, especially for the missing transverse energy. However, as the event yield seems to increase for higher $E_{\mathrm{T}}^{\mathrm{miss}}$, the change caused by the bug correction might even slightly improve the sensitivity for lower stau masses in the highMass signal region ($E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV).

To further assess the possible impact on the sensitivity of the analysis, a comparison between the two software releases has been also performed using the signal region selections of the first iteration of the stau search at truth level. The reason for not using the BDT-based signal regions of the second analysis iteration lies in the technical effort of preparing a sample at truth-level such that the BDTs can be evaluated on it. Moreover, the cut-based regions are sufficient for a first check of the potential change in event yields in the most sensitive regions of phase-space. The corresponding plots can be found in Fig. G.5. It has to be noted that the shown event yields cannot be directly compared to the number of events after the full reconstruction process as the latter introduces a selection efficiency that has to be considered as well.

The event yields for a stau mass of 360 GeV are in reasonable agreement. However, there are some deviations for a stau mass of 120 GeV. While there are very large statistical uncertainties for $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 60)$ GeV, the sample for the $(120, 1)$ GeV model was produced with enhanced statistics particularly for this purpose. From its event yields in the lowMass region a clear decrease of 20% can be observed for the software release that incorporates the bug correction. Potentially, this could cause a degradation of the sensitivity for stau pair production models with low stau masses, compared to the original expectation. Again using the example of the lowMass SR from the first analysis iteration, the behaviour of the significance in dependence on the number of signal events is shown

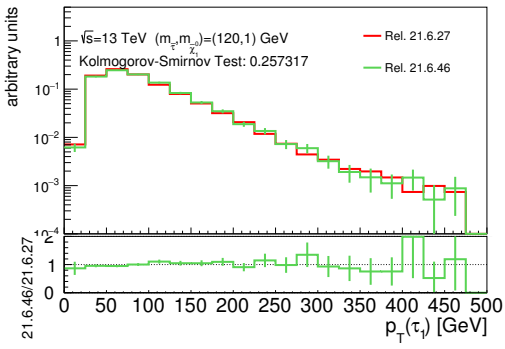
in Fig. G.6. In the same figure the effect on the significance through a 20% decrease in signal yields is illustrated as well. Although the drop in the Z-value is larger for a larger number of signal events, the significance curves in dependence of the signal yield are not very steep. A decrease to significances below the critical value of 1.64 is only possible for significances that have already been narrowly above this threshold before the bug correction. The typical significances in the lowMass signal region for the second analysis iteration are at 2.2 to 3.2, which could mean that the sensitivity will not be decreased below exclusion sensitivity for most of the models in the lowMass regime. A comparison of the significances per signal model with a 20% decrease of signal events with respect to the original expected exclusion sensitivity for the BDT-based and binned SR-lowMass is shown in Fig. G.7. Although the area with exclusion sensitivity becomes smaller, there is still sufficient sensitivity for exclusion for most of the models with low stau masses. (Starting at a stau mass of $\gtrsim 180\,\text{GeV}$, the largest contribution to the overall significance will originate from SR-highMass.)

However, these statements rely on various assumptions: Firstly, the 20% yield decrease for the $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$ should be a good approximation for other models in the low stau mass regime. Secondly, the decrease in events observed at truth level should translate into a similar decrease at reconstruction level. And thirdly, the results of the checks in the cut-based lowMass signal region should (approximately) hold in the BDT-based lowMass signal region as well. As none of these assumptions necessarily has to be true, the full impact on the exclusion sensitivity can only be assessed as soon as the simulations of all the signal models have been reproduced and evaluated at reconstruction-level in the BDT-based signal regions defined in Tab. 7.3 and Tab. 7.4.

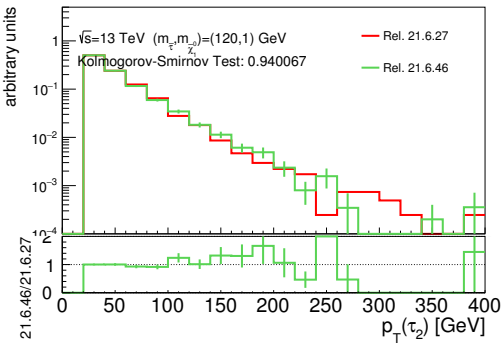


(a) $E_{\mathrm{T}}^{\mathrm{miss}}$

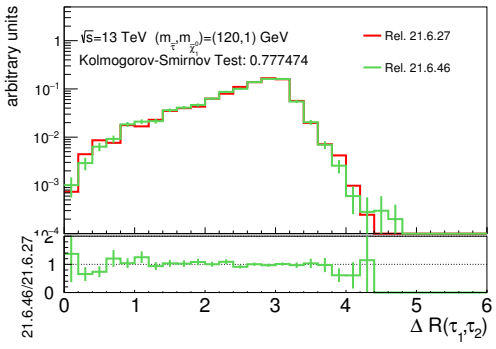
(b) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



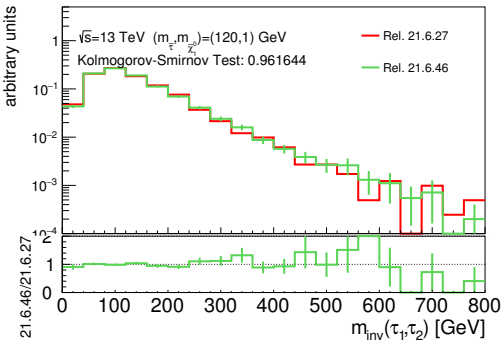
(c) $p_{\mathrm{T}}(\tau_1)$



(d) $p_{\mathrm{T}}(\tau_2)$

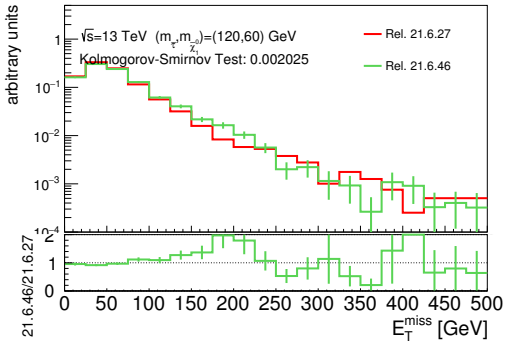


(e) $\Delta R(\tau_1, \tau_2)$

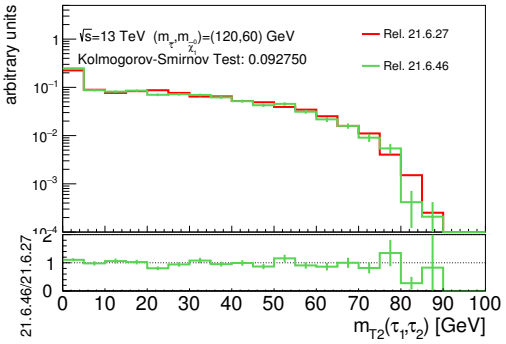


(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

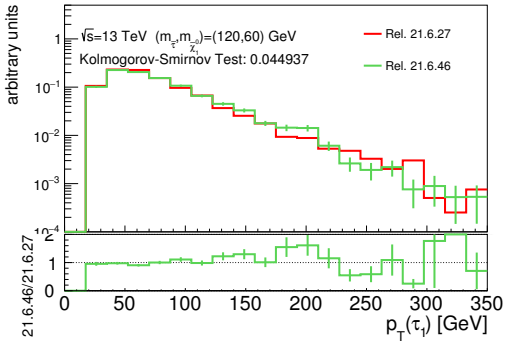
Figure G.1.: Comparison of simulation software versions for $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\mathrm{GeV}$.



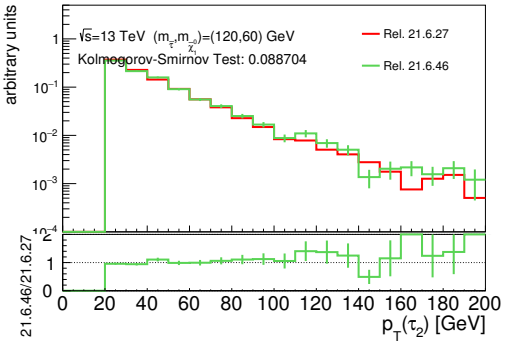
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$



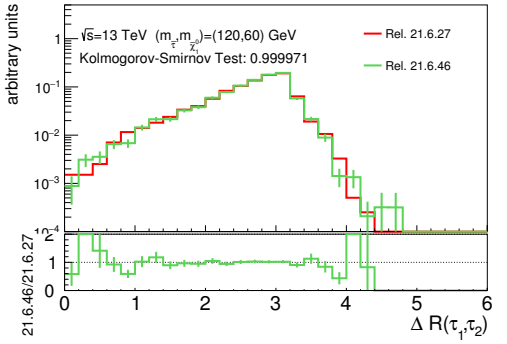
(b) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



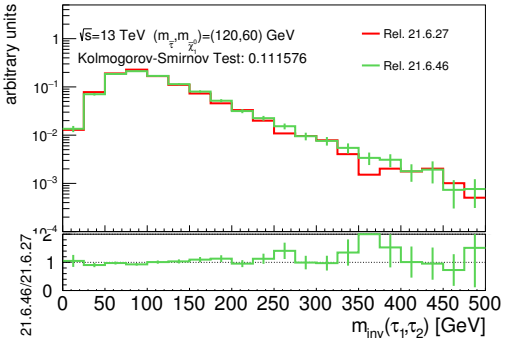
(c) $p_{\mathrm{T}}(\tau_1)$



(d) $p_{\mathrm{T}}(\tau_2)$

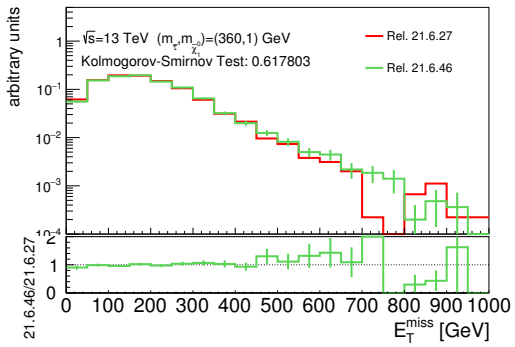


(e) $\Delta R(\tau_1, \tau_2)$

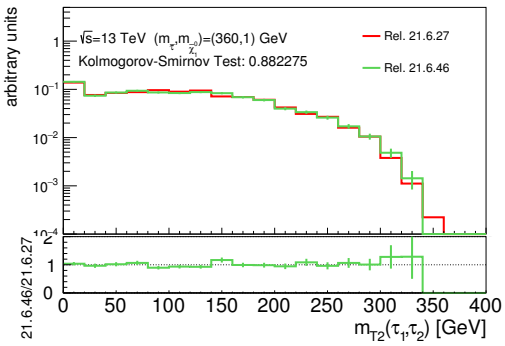


(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

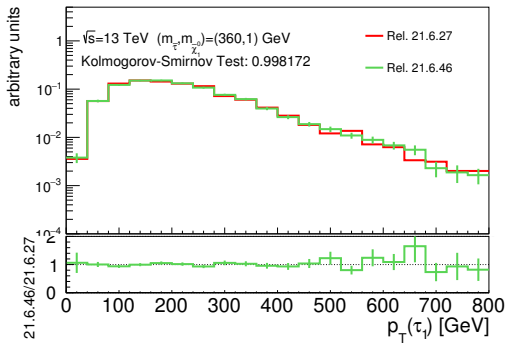
Figure G.2.: Comparison of simulation software versions for $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 60)$ GeV.



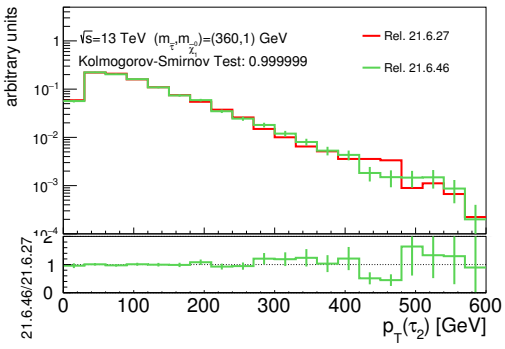
(a) $E_{\mathrm{T}}^{\mathrm{miss}}$



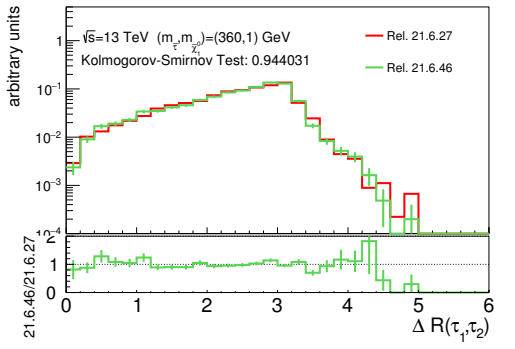
(b) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



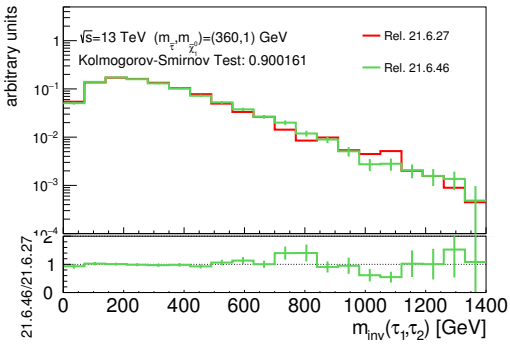
(c) $p_{\mathrm{T}}(\tau_1)$



(d) $p_{\mathrm{T}}(\tau_2)$

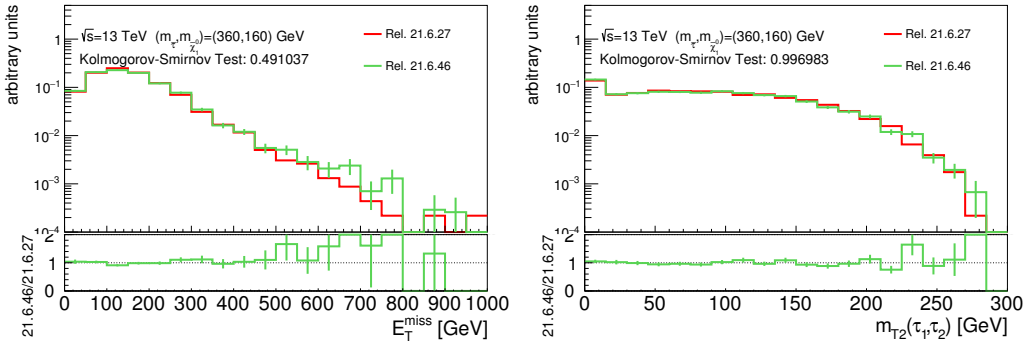


(e) $\Delta R(\tau_1, \tau_2)$



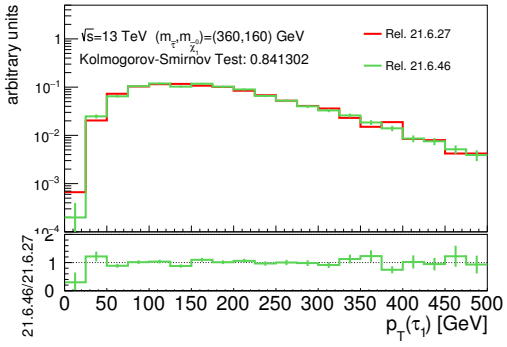
(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

Figure G.3.: Comparison of simulation software versions for $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)\,\mathrm{GeV}$.

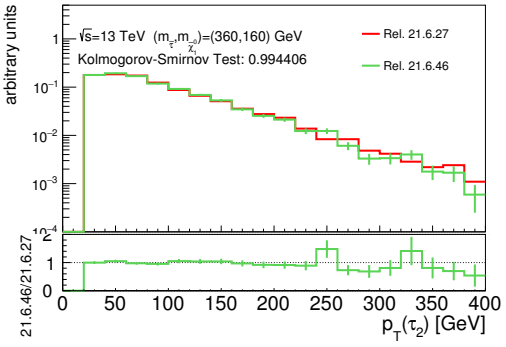


(a) $E_{\mathrm{T}}^{\mathrm{miss}}$

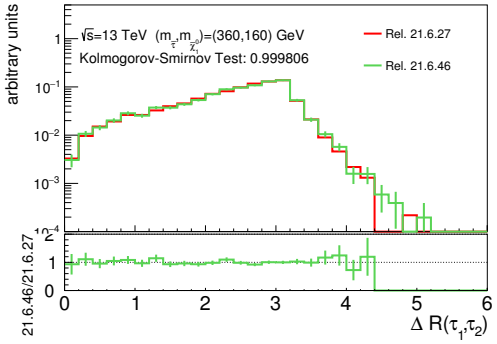
(b) $m_{\mathrm{T2}}(\tau_1, \tau_2)$



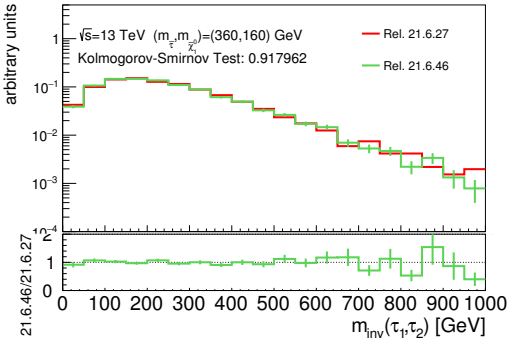
(c) $p_{\mathrm{T}}(\tau_1)$



(d) $p_{\mathrm{T}}(\tau_2)$

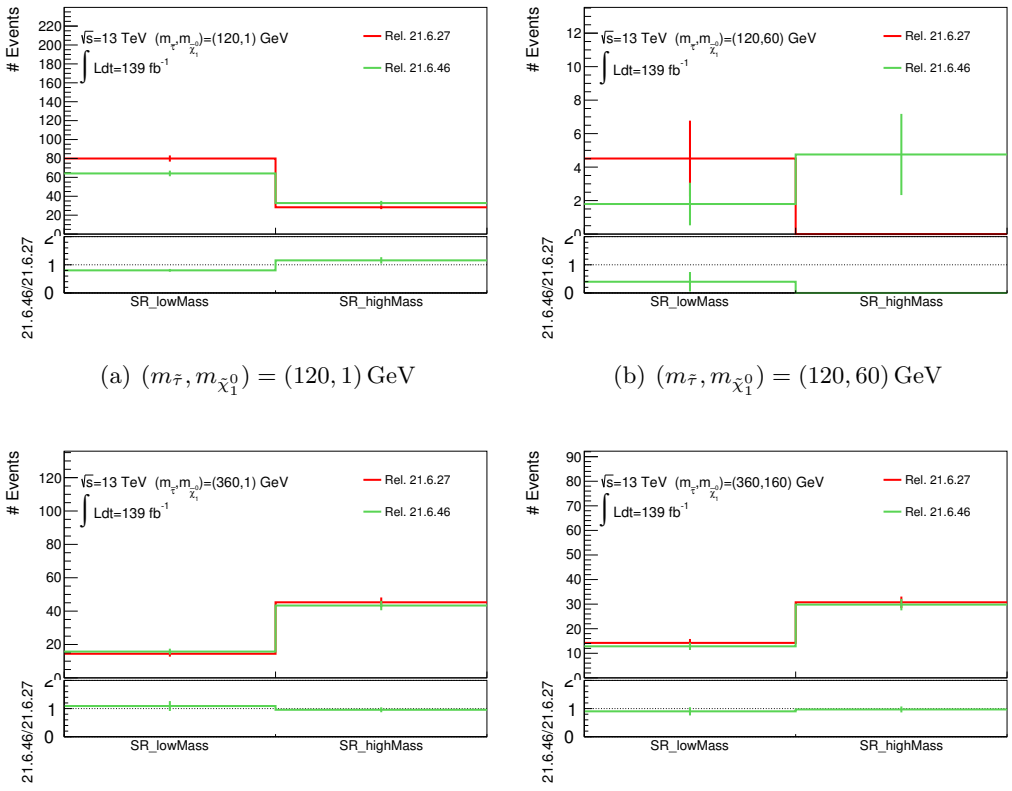


(e) $\Delta R(\tau_1, \tau_2)$



(f) $m_{\mathrm{inv}}(\tau_1, \tau_2)$

Figure G.4.: Comparison of simulation software versions for $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 160)$ GeV.



(a) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$

(b) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 60)\,\text{GeV}$

(c) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 1)\,\text{GeV}$

(d) $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (360, 160)\,\text{GeV}$

Figure G.5.: Comparison of simulation software versions in the signal regions defined in Tab.6.3. The left (right) bin shows the event yield for SR-lowMass (SR-highMass). While there is negligible change for the high stau mass models, the event yield for the low stau mass models decreases in SR-lowMass with small increase in SR-highMass. This could cause a deterioration of the sensitivity for low stau masses.

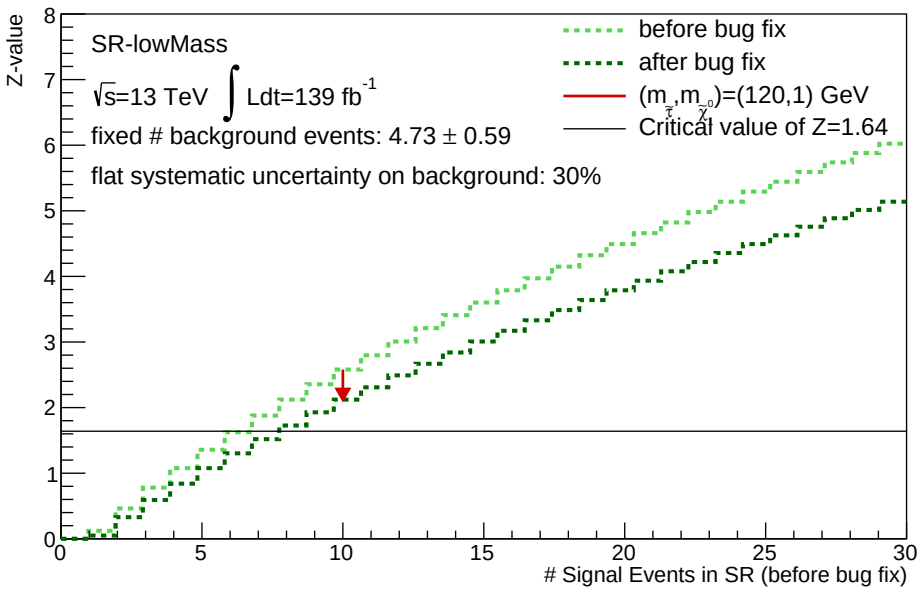
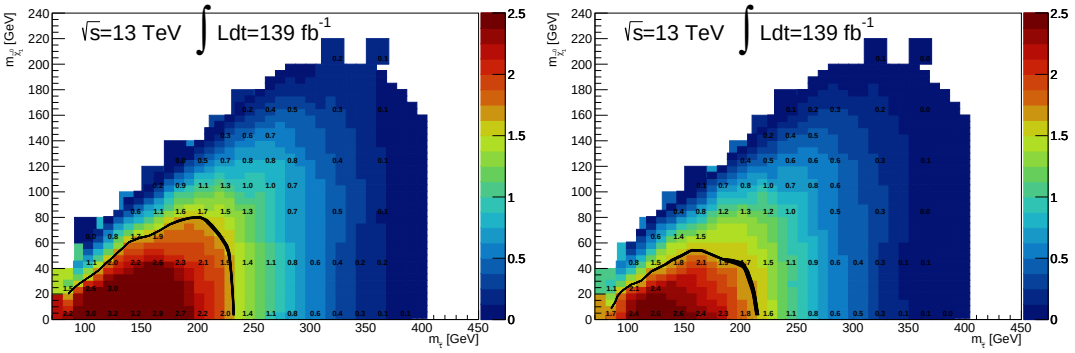


Figure G.6.: Dependence of the significance in SR-lowMass on the signal yield. The x-axis shows the number of possible signal events in SR-lowMass. The light green line corresponds to the significance value Z calculated with the signal yield values from the x-axis and a fixed background yield. The dark green line corresponds to the significance resulting in the case of a decrease in signal yields by 20% of the respective x-axis value. Therefore, the change due to the bug correction is approximately equivalent to a vertical shift from the light green line to the dark green line. The particular case for the signal model $(m_{\tilde{\tau}}, m_{\tilde{\chi}_1^0}) = (120, 1)\,\text{GeV}$ is indicated by the red arrow. Although the significance decreases, the Z-value is still above the critical value of 1.64.



(a) Full signal yield

(b) -20% of signal yield

Figure G.7.: Change of exclusion sensitivity for a signal yield reduction of 20%. The left plot is the same as shown in Fig. 7.4. For the right plot only 80% of the signal yields have been taken into account. For both plots, a flat systematic uncertainty of 30% on the background yield and the statistical background uncertainty have been taken into account.

Bibliography

[1] Particle Data Group. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/33653179/ptaa104.pdf.

[2] D. Schaile. *Lecture Notes on "Advanced Particle Physics"*. LMU München, Germany. 2020.

[3] M. Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 9781107034266.

[4] S. Tomonaga. "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields". In: *Progress of Theoretical Physics* 1.2 (1946), pp. 27–42. DOI: 10.1143/PTP.1.27. eprint: /oup/backfile/content_public/journal/ptp/1/2/10.1143/ptp.1.27/2/1-2-27.pdf.

[5] Julian Schwinger. "On Quantum-Electrodynamics and the Magnetic Moment of the Electron". In: *Phys. Rev.* 73 (4 1948), pp. 416–417. DOI: 10.1103/PhysRev.73.416.

[6] J. Schwinger. "Quantum Electrodynamics. I. A Covariant Formulation". In: *Physical Review* 74 (1948), pp. 1439–1461. DOI: 10.1103/PhysRev.74.1439.

[7] R. P. Feynman. "Space-Time Approach to Quantum Electrodynamics". In: *Phys. Rev.* 76 (6 1949), pp. 769–789. DOI: 10.1103/PhysRev.76.769.

[8] R. P. Feynman. "Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction". In: *Physical Review* 80 (1950), pp. 440–457. DOI: 10.1103/PhysRev.80.440.

[9] D. J. Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008.

[10] Michio Kaku. *Quantum Field Theory - A Modern Introduction*. New York: Oxford University Press, 1993. ISBN: 9780195076523.

[11] D.V.Schroeder M.E.Peskin. *An Introduction to Quantum Field Theory*. CRC Press, 1995. ISBN: 9780201503975.

[12] A. Bednyakov. *Lecture Notes on "The E-W Standard Model"*. European School of High-Energy Physics 2019, St. Petersburg, Russia. 2019.

[13] Enrico Fermi. "Versuch einer Theorie der β-Strahlen". In: *Zeitschrift für Physik* 88 (1934), pp. 161–177. DOI: 10.1007/BF01351864.

[14] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413.

[15] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[16] Makoto Kobayashi and Toshihide Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf.

[17] P. Schmüser. *Feynman-Graphen und Eichtheorien für Experimentalphysiker*. Berlin Heidelberg New York: Springer-Verlag, 1995. ISBN: 9783540584865.

[18] Sheldon L. Glashow. “The renormalizability of vector meson interactions”. In: *Nuclear Physics* 10 (1959), pp. 107–117. DOI: 10.1016/0029-5582(59)90196-8.

[19] Abdus Salam. “Weak and electromagnetic interactions”. In: *Il Nuovo Cimento* 11 (1959), pp. 568–577. DOI: 10.1007/BF02726525.

[20] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[21] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[22] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[23] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[24] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (). ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020.

[25] CMS Collaboration. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[26] Yoichiro Nambu. “Quasi-Particles and Gauge Invariance in the Theory of Superconductivity”. In: *Phys. Rev.* 117 (3 1960), pp. 648–663. DOI: 10.1103/PhysRev.117.648.

[27] J. Goldstone. “Field theories with “superconductor” solutions”. In: *Il Nuovo Cimento* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722.

[28] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. “Broken Symmetries”. In: *Phys. Rev.* 127 (3 1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965.

[29] Y. Fukuda et al. “Evidence for Oscillation of Atmospheric Neutrinos”. In: *Physical Review Letters* 81.8 (1998), 1562–1567. ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562.

[30] Q. R. Ahmad et al. “Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory”. In: *Physical Review Letters* 87.7 (2001). ISSN: 1079-7114. DOI: 10.1103/physrevlett.87.071301. URL: http://dx.doi.org/10.1103/PhysRevLett.87.071301.

[31] Ugo Amaldi, Wim de Boer, and H Fürstenau. “Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP”. In: *Phys. Lett. B* 260 (1991), pp. 447–455. DOI: 10.1016/0370-2693(91)91641-8. URL: http://cds.cern.ch/record/217892.

[32] D.I. Kazakov. "Beyond the standard model: In search of supersymmetry". In: *2000 European School of High-Energy Physics*. Aug. 2000, pp. 125–199. arXiv: hep-ph/0012288.

[33] Stephen P. Martin. "A Supersymmetry primer". In: (). [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. DOI: 10.1142/9789812839657_0001, 10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[34] V. Rubakov. *Lecture Notes on "Cosmology and Dark Matter"*. European School of High-Energy Physics 2010, St. Petersburg, Russia. 2019.

[35] Sidney Coleman and Jeffrey Mandula. "All Possible Symmetries of the S Matrix". In: *Phys. Rev.* 159 (1967), pp. 1251–1256. DOI: 10.1103/PhysRev.159.1251. URL: https://link.aps.org/doi/10.1103/PhysRev.159.1251.

[36] Hironari Miyazawa. "Spinor Currents and Symmetries of Baryons and Mesons". In: *Phys. Rev.* 170 (5 1968), pp. 1586–1590. DOI: 10.1103/PhysRev.170.1586. URL: https://link.aps.org/doi/10.1103/PhysRev.170.1586.

[37] Melissa van Beekveld et al. "Supersymmetry with Dark Matter is still natural". In: *Phys. Rev.* D96.3 (2017), p. 035015. DOI: 10.1103/PhysRevD.96.035015. arXiv: 1612.06333 [hep-ph].

[38] Daniel Albornoz Vasquez, Genevieve Belanger, and Celine Boehm. "Revisiting light neutralino scenarios in the MSSM". In: *Phys. Rev. D* 84 (2011), p. 095015. DOI: 10.1103/PhysRevD.84.095015. arXiv: 1108.1338 [hep-ph].

[39] Alexander Mann et al. URL: https://github.com/fuenfundachtzig/xsec, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections.

[40] G. Bozzi, B. Fuks, and M. Klasen. "Transverse-momentum resummation for slepton-pair production at the CERN Large Hadron Collider". In: *Phys. Rev. D* 74 (2006), p. 015001. DOI: 10.1103/PhysRevD.74.015001. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.015001.

[41] Benjamin Fuks et al. "Precision predictions for electroweak superpartner production at hadron colliders with Resummino". In: *The European Physical Journal C* 73.2480 (2013). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2480-0. URL: http://dx.doi.org/10.1140/epjc/s10052-013-2480-0.

[42] Benjamin Fuks et al. "Revisiting slepton pair production at the Large Hadron Collider". In: *Journal of High Energy Physics* 2014.1 (2014). ISSN: 1029-8479. DOI: 10.1007/jhep01(2014)168. URL: http://dx.doi.org/10.1007/JHEP01(2014)168.

[43] Juri Fiaschi and Michael Klasen. "Slepton pair production at the LHC in NLO+NLL with resummation-improved parton densities". In: *Journal of High Energy Physics* 2018.3 (2018). ISSN: 1029-8479. DOI: 10.1007/jhep03(2018)094. URL: http://dx.doi.org/10.1007/JHEP03(2018)094.

[44] W. Beenakker et al. "Production of Charginos, Neutralinos, and Sleptons at Hadron Colliders". In: *Phys. Rev. Lett.* 83 (19 1999), pp. 3780–3783. DOI: 10.1103/PhysRevLett.83.3780. URL: https://link.aps.org/doi/10.1103/PhysRevLett.83.3780.

[45] Clara E. Leitgeb. "Search for Direct Production of Supersymmetric Scalar Tau Leptons Using 13 TeV Data Taken with the ATLAS Detector at the LHC". MA thesis. Ludwig-Maximilians Universität München, 2017 (not public).

[46] ATLAS SUSY Group. *SUSY Feynman Diagrams*. (not public). URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SUSYFeynmanDiagrams.

[47] LEPSUSYWG et al. *note LEPSUSYWG/04-01.1*. URL: http://lepsusy.web.cern.ch/lepsusy/Welcome.html.

[48] ATLAS Collaboration. "Search for the direct production of charginos, neutralinos and staus in final states with at least two hadronically decaying taus and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 10 (2014), p. 096. DOI: 10.1007/JHEP10(2014)096. arXiv: 1407.0350 [hep-ex].

[49] CMS Collaboration. *Search for pair production of tau sleptons in $\sqrt{s} = 13$ TeV pp collisions in the all-hadronic final state*. Tech. rep. CMS-PAS-SUS-17-003. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2273395.

[50] CMS Collaboration. "Search for direct pair production of supersymmetric partners to the τ lepton in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *The European Physical Journal C* 80.189 (2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-7739-7. URL: http://dx.doi.org/10.1140/epjc/s10052-020-7739-7.

[51] ATLAS Collaboration. "Search for direct stau production in events with two hadronic τ-leptons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *Physical Review D* 101.3 (2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.101.032009. URL: http://dx.doi.org/10.1103/PhysRevD.101.032009.

[52] *Linear accelerator 2*. URL: https://cds.cern.ch/record/1997427.

[53] *The Proton Synchrotron Booster*. URL: https://cds.cern.ch/record/1997372.

[54] *The Proton Synchrotron*. URL: https://cds.cern.ch/record/1997189.

[55] *The Super Proton Synchrotron*. URL: https://cds.cern.ch/record/1997188.

[56] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[57] *LHC Facts*. URL: http://www.lhc-facts.ch/.

[58] E Senes et al. "Transverse emittance measurement in the CERN Proton Synchrotron in view of beam production for the High-Luminosity LHC". In: *Journal of Physics: Conference Series* 1350 (2019), p. 012087. DOI: 10.1088/1742-6596/1350/1/012087.

[59] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[60] CMS Collaboration. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[61] LHCb Collaboration. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[62] ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[63] Joao Pequenao 2015 (2017). *CERN Document Server*. URL: https://cds.cern.ch/images/CERN-GE-0803012-01.

[64] Joao Pequenao. "Computer generated image of the ATLAS inner detector". 2008. URL: https://cds.cern.ch/record/1095926.

[65] Joao Pequenao. "Computer Generated image of the ATLAS calorimeter". 2008. URL: https://cds.cern.ch/record/1095927.

[66] Joao Pequenao. "Computer generated image of the ATLAS Muons subsystem". 2008. URL: https://cds.cern.ch/record/1095929.

[67] Aranzazu Ruiz-Martinez and ATLAS Collaboration. *The Run-2 ATLAS Trigger System*. Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, 2016. DOI: 10.1088/1742-6596/762/1/012003. URL: https://cds.cern.ch/record/2133909.

[68] M. Abolins et al. "The ATLAS Data Acquisition and High Level Trigger system". In: *JINST* 11.06 (2016), P06008. DOI: 10.1088/1748-0221/11/06/P06008.

[69] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, 2019. URL: http://cds.cern.ch/record/2677054.

[70] G.Cowan. *Statistical Data Analysis*. Oxford New York: Oxford Univerity Press, 1998. ISBN: 0198501560.

[71] Richard D. Ball et al. "Parton distributions for the LHC run II". In: *Journal of High Energy Physics* 2015.4 (2015). ISSN: 1029-8479. DOI: 10.1007/jhep04(2015)040. URL: http://dx.doi.org/10.1007/JHEP04(2015)040.

[72] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nuclear Physics B* 867.2 (2013), 244–289. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2012.10.003. URL: http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003.

[73] A. D. Martin et al. "Parton distributions for the LHC". In: *The European Physical Journal C* 63.2 (2009), 189–285. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-009-1072-5. URL: http://dx.doi.org/10.1140/epjc/s10052-009-1072-5.

[74] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *The European Physical Journal C* 70.3 (2010), 823–874. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1429-9. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1429-9.

[75] Torbjorn Sjostrand. "Monte Carlo Generators". In: *High-energy physics. Proceedings, European School, Aronsborg, Sweden, June 18-July 1, 2006*. 2006, pp. 51–74. arXiv: hep-ph/0611247 [hep-ph]. URL: http://weblib.cern.ch/abstract?CERN-LCGAPP-2006-06.

[76] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[77] B. Webber. "Parton shower Monte Carlo event generators". In: *Scholarpedia* 6.12 (2011). revision #128236, p. 10662. DOI: 10.4249/scholarpedia.10662.

[78] S. Agostinelli et al. "GEANT4–a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[79] Judith M. Katzy. "QCD Monte-Carlo model tunes for the LHC". In: *Progress in Particle and Nuclear Physics* 73 (2013), pp. 141 –187. ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2013.08.002. URL: http://www.sciencedirect.com/science/article/pii/S0146641013000768.

[80] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.12 (2019), P12006–P12006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/12/p12006. URL: http://dx.doi.org/10.1088/1748-0221/14/12/P12006.

[81] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1449796.

[82] ATLAS Collaboration. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s}$ = 13 TeV". In: *The European Physical Journal C* 79.8 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7140-6.

[83] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.03 (2019), P03017–P03017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/03/p03017. URL: http://dx.doi.org/10.1088/1748-0221/14/03/P03017.

[84] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *The European Physical Journal C* 76.5 (2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y.

[85] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[86] ATLAS Collaboration. *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. 2020. arXiv: 2007.02645 [hep-ex].

[87] ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. ATL-PHYS-PUB-2017-013. Geneva: CERN, 2017. URL: http://cds.cern.ch/record/2273281.

[88] ATLAS Collaboration. *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. ATLAS-CONF-2017-029. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2261772.

[89] ATLAS Collaboration. *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. ATL-PHYS-PUB-2015-045. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2064383.

[90] Clara E. Leitgeb. *Performance Studies of Electron-Tau Discriminators via Tag-and-Probe Measurements in $Z \to ee$ Events and Scale-Factor Calculation in 2017 ATLAS Data*. Tech. rep. ATL-COM-PHYS-2019-115. Geneva: CERN, 2019 (not public). URL: https://cds.cern.ch/record/2659602.

[91] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*. 2018. arXiv: 1802.08168 [hep-ex].

[92] C. G. Lester and D. J. Summers. "Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders". In: *Phys. Lett.* B463 (1999), pp. 99–103. DOI: 10.1016/S0370-2693(99)00945-4. arXiv: hep-ph/9906349 [hep-ph].

[93] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[94] Robert D. Cousins, James T. Linnemann, and Jordan Tucker. "Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 595.2 (2008), 480–501. ISSN: 0168-9002. DOI: 10.1016/j.nima.2008.07.086. URL: http://dx.doi.org/10.1016/j.nima.2008.07.086.

[95] A L Read. "Presentation of search results: the CLs technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313. URL: https://doi.org/10.1088%2F0954-3899%2F28%2F10%2F313.

[96] M. Baak et al. "HistFitter software framework for statistical data analysis". In: *The European Physical Journal C* 75.4 (2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3327-7. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3327-7.

[97] Enrico Bothmann et al. "Event Generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[98] T Gleisberg et al. "Event generation with SHERPA 1.1". In: *Journal of High Energy Physics* 2009.02 (2009), pp. 007–007. DOI: 10.1088/1126-6708/2009/02/007. URL: https://doi.org/10.1088/1126-6708/2009/02/007.

[99] Stefan Höche et al. "QCD matrix elements + parton showers. The NLO case". In: *Journal of High Energy Physics* 2013.4 (2013). ISSN: 1029-8479. DOI: 10.1007/jhep04(2013)027. URL: http://dx.doi.org/10.1007/JHEP04(2013)027.

[100] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". In: *Journal of High Energy Physics* 2008.12 (2008), pp. 039–039. DOI: 10.1088/1126-6708/2008/12/039. URL: https://doi.org/10.1088/1126-6708/2008/12/039.

[101] F. Cascioli, P. Maierhöfer, and S. Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (11 2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.111601.

[102] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. "Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations". In: *Comput. Phys. Commun.* 212 (2017), pp. 220–238. DOI: 10.1016/j.cpc.2016.10.013. arXiv: 1604.06792 [hep-ph].

[103] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[104] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[105] Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. “A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *Journal of High Energy Physics* 2007.09 (2007), pp. 126–126. DOI: 10.1088/1126-6708/2007/09/126.

[106] Paolo Nason. “A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms”. In: *Journal of High Energy Physics* 2004.11 (2004), pp. 040–040. DOI: 10.1088/1126-6708/2004/11/040.

[107] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. “A brief introduction to PYTHIA 8.1”. In: *Computer Physics Communications* 178.11 (2008), 852–867. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2008.01.036.

[108] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *Journal of High Energy Physics* 2014.7 (2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079.

[109] D. de Florian et al. “Handbook of LHC higgs cross sections: 4. Deciphering the nature of the Higgs sector”. In: (2017). DOI: 10.23731/CYRM-2017-002.

[110] Lorenzo Moneta et al. “The RooStats Project”. In: *PoS* ACAT2010 (2010), p. 057. arXiv: 1009.1003 [physics.data-an].

[111] R. Brun and F. Rademakers. “ROOT: An object oriented data analysis framework”. In: *Nucl. Instrum. Meth. A* 389 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.

[112] Daniel R Tovey. “On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders”. In: *Journal of High Energy Physics* 2008.04 (2008), 034–034. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/034. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/034.

[113] ATLAS Collaboration. *Improved description of the di-tau final state in events with associated production of a W boson and jets in the ATLAS detector using the tau-promotion method*. Tech. rep. ATL-PHYS-PUB-2019-039. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2692073.

[114] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: *Journal of Physics G: Nuclear and Particle Physics* 43.2 (2016), p. 023001. ISSN: 1361-6471. DOI: 10.1088/0954-3899/43/2/023001. URL: http://dx.doi.org/10.1088/0954-3899/43/2/023001.

[115] Manuel Bähr et al. “Herwig++ physics and manual”. In: *The European Physical Journal C* 58.4 (2008), 639–707. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-008-0798-9. URL: http://dx.doi.org/10.1140/epjc/s10052-008-0798-9.

[116] Robert E. Schapire. "The Strength of Weak Learnability". In: *Mach. Learn.* 5 (1990), pp. 197–227. DOI: 10.1007/BF00116037. URL: https://doi.org/10.1007/BF00116037.

[117] Yoav Freund and Robert E Schapire. "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting". In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119 –139. ISSN: 0022-0000. DOI: http://dx.doi.org/10.1006/jcss.1997.1504. URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[118] Robert E. Schapire. *Explaining AdaBoost*. Jan. 2013, pp. 37–52. ISBN: 978-3-642-41135-9.

[119] Leo Breiman. "Prediction Games and Arcing Algorithms". In: *Neural Computation* 11.7 (1999), pp. 1493–1517. DOI: 10.1162/089976699300016106. URL: https://doi.org/10.1162/089976699300016106.

[120] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." In: *Ann. Statist.* 29.5 (Oct. 2001), pp. 1189–1232. DOI: 10.1214/aos/1013203451. URL: https://doi.org/10.1214/aos/1013203451.

[121] L. Breiman. "Arcing the edge". In: *Ann. Prob.* 26 (1998), pp. 1683–1702.

[122] A. N. Kolmogorov. "Sulla Determinazione Empirica di una Legge di Distribuzione". In: *Giornale dell'Istituto Italiano degli Attuari* 4 (1933), pp. 83–91.

[123] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[124] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn & TensorFlow*. Sebastopol: O'Reilly, 2017. ISBN: 9781491962299.

[125] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[126] ATLAS Collaboration. *Prospects for searches for staus, charginos and neutralinos at the high luminosity LHC with the ATLAS Detector*. Tech. rep. ATL-PHYS-PUB-2018-048. Geneva: CERN, 2018. URL: http://cds.cern.ch/record/2651927.

List of Figures

List of Tables